Computer Science & Information Technology 176

Artificial Intelligence and Fuzzy Logic System

David C. Wyld,
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

- 8th International Conference on Artificial Intelligence and Fuzzy Logic System (AIFZ 2022)
- 8th International Conference on Signal and Image Processing (SIGPRO 2022)
- 8th International Conference on Networks & Communications (NWCOM 2022)
- 3rd International Conference on Cloud Computing, Security and Blockchain (CLSB 2022)
- 3rd International Conference on Machine Learning Techniques and NLP (MLNLP 2022)
- 8th International Conference on Data Mining (DTMN 2022)
- 3rd International Conference on Big Data & IoT (BDIoT 2022)
- 8th International Conference on Computer Science, Engineering and Information Technology (CSITY 2022)

Published By

AIRCC Publishing Corporation

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

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

ISSN: 2231 - 5403
ISBN: 978-1-925953-76-3
DOI: 10.5121/csit.2022.121601 - 10.5121/csit.2022.121616

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

Preface

8th International Conference on Artificial Intelligence and Fuzzy Logic System (AIFZ 2022), 8th International Conference on Signal and Image Processing (SIGPRO 2022), 8th International Conference on Networks & Communications (NWCOM 2022), 3rd International Conference on Cloud Computing, Security and Blockchain (CLSB 2022), 3rd International Conference on Machine Learning Techniques and NLP (MLNLP 2022), 8th International Conference on Data Mining (DTMN 2022), 3rd International Conference on Big Data & IoT (BDIoT 2022), 8th International Conference on Computer Science, Engineering and Information Technology (CSITY 2022) was collocated with 8th International Conference on Artificial Intelligence and Fuzzy Logic System (AIFZ 2022). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

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

The AIFZ 2022, SIGPRO 2022, NWCOM 2022, CLSB 2022, MLNLP 2022, DTMN 2022, BDIOT 2022 and CSITY 2022. Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically.

In closing, AIFZ 2022, SIGPRO 2022, NWCOM 2022, CLSB 2022, MLNLP 2022, DTMN 2022, BDIOT 2022 and CSITY 2022 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the AIFZ 2022, SIGPRO 2022, NWCOM 2022, CLSB 2022, MLNLP 2022, DTMN 2022, BDIOT 2022 and CSITY 2022.

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

David C. Wyld,
Dhinaharan Nagamalai (Eds)

General Chair	Organization
David C. Wyld,	Southeastern Louisiana University, USA
Dhinaharan Nagamalai (Eds)	Wireilla Net Solutions, Australia

Program Committee Members

Abbas Khosravi,	Deakin University, Australia
Abdel-Badeeh M. Salem,	Ain Shams University, Egypt
Abdelhadi Assir,	Hassan 1st University, Morocco
Abdellatif I. Moustafa,	Umm AL-Qura University, Saudi Arabia
Abderrahim Siam,	University of Khenchela, Algeria
Abderrahmane Ez-zahout,	Mohammed V University, Morocco
Abhishek Shukla,	R D Engineering College, India
Addisson Salazar,	Universitat Politècnica de València, Spain
Adnan M. Mussein,	Northern Technical University, Iraq
Adrian Olaru,	University Politehnica of Bucharest, Romania
Ahmad A. Saifan,	Yarmouk University, Jordan
Ahmad Alazzam,	Jadara University, Jordan
Ahmed Yaseen Mjhool,	University of Kufa, Iraq
Akhil Gupta,	Lovely Professional University, India
Alessio Ishizaka,	NEOMA Business School, France
Alex Mathew,	Bethany College, USA
Ali Al-Sabbagh,	Ministry of communication, Iraq
Ali Asif,	Harbin Engineering University, China
Alireza Valipour Baboli,	Technical and Vocational University, Iran
Amari Houda,	Networking & Telecom Engineering, Tunisia
Amel Ourici,	Badji Mokhtar University of Annaba, Algeria
Amin Karami,	University of East London (UEL), UK
Amizah Malip,	University of Malaya, Malaysia
Ana Luísa Varani Leal,	University of Macau, Macau
Anand J Kulkarni,	MIT World Peace University, India
Anand Nayyar,	Duy Tan University, Vietnam
Anand Singh Rajawat,	Sandip University, India
Andy Rachman,	Institut Teknologi Adhi Tama Surabaya, Indonesia
Anouar Abtoy,	Abdelmalek Essaâdi University, Morocco
António Abreu,	ISEL, Portugal
Aridj mohamed,	University Hassiba Benbouali Chlef, Algeria
Arighna Basak,	Brainware University, India
Armir Bujaria,	Studi di Padova, Italy
Arnold Kwofie,	University for Development Studies, Ghana
Assia Djenouhat,	University Algiers 3, Algeria
Atul Garg,	Chitkara University, India
Bello Abdulazeez,	Ignatius Ajuru University, Nigeria
Benyettou Mohammed,	University center of Relizane, Algeria
Beshair Alsiddiq,	Prince Sultan University, Saudi Arabia
Bin Zhao,	JD.com Silicon Valley R&D Center, USA
Bouchra Marzak,	Hassan II University, Morocco
Brahim Lejdel,	University of El-Oued, Algeria
Brij Gupta,	National Institute of Technology Kurukshetra, India
Cagdas Hakan Aladag,	Hacettepe University, Turkey

Carla Osthoff,	National Laboratory for Scientific Computing, Brazil
Carlos Becker Westphall,	Federal University of Santa Catarina, Brazil
Cheng Siong Chin,	Newcastle University, Singapore
Cherkaoui Leghris,	Hassan II University of Casablanca, Morocco
Chin-Chen Chang,	Feng Chia University, Taiwan
Chittineni Suneetha,	R.V.R & J.C. College of Engineering, India
Christian Mancas,	Ovidius University, Romania
Dadmehr Rahbari,	University Of Qom, Iran
Dan Wan,	Hunan Normal University, China
Daniel Hunyadi,	"Lucian Blaga" University of Sibiu, Romania
Dário Ferreira,	Universidade da Beira Interior, Portugal
Dmitry G. Korzun,	Petrozavodsk State University, Russia
Domenico Rotondi,	Fincons SpA, Italy
El Habib Nfaoui,	Sidi Mohamed Ben Abdellah University, Morocco
EL Murabet Amina,	Abdelmalek Essaadi University, Morocco
Elżbieta Macioszek,	Silesian University of Technology, Poland
Emir Kremic,	Federal Institute of Statistics, Bosnia and Herzegovina
Eugénia Moreira Bernardino,	Polytechnic Institute of Leiria, Portugal
F. Abbasi,	Islamic Azad University, Iran
Faouzia Benabbou,	University Hassan II of Casablanca, Morocco
Felix J. Garcia Clemente,	University of Murcia, Spain
Fernando Zacarias Flores,	Universidad Autonoma de Puebla, Mexico
Fitri Utaminingrum,	Brawijaya University, Indonesia
Francesco Zirilli,	Sapienza Universita Roma, Italy
Fzlollah Abbasi,	Islamic Azad University, Iran
Gajendra Sharma,	Kathmandu University, Nepal
Gladston Raj S,	Govt. College, India
Grigorios N. Beligiannis,	University of Patras - Agrinio Campus, Greece
Grzegorz Sierpiński,	Silesian University of Technology, Poland
Gueddah Hicham,	Mohammed V University, Morocco
Guezouli Larbi,	Higher National School of Renewable Energies, Algeria
Gülden Köktürk,	Dokuz Eylül University, Turkey
Gunasekaran Perumalsamy,	Ramco Institute of Technology, India
Gyu Myoung Lee,	Liverpool John Moores University, UK
Hala Abukhalaf,	Palestine Polytechnic University, Palestine
Hamed Taherdoost,	West University, Canada
Hamid Ali Abed AL-Asadi,	Iraq University College, Iraq
Hamid Khemissa,	USTHB University Algiers, Algeria
Hedayat Omidvar,	National Iranian Gas Company, Iran
Hicham Gueddah,	Mohammed V University in Rabat, Morocco
Hlaing Htake Khaung Tin,	University of Information Technology, Myanmar
Ilham Huseyinov,	Istanbul Aydin University, Turkey
Ines Bayoudh Saadi,	Tunis University, Tunisia
Isa Maleki,	Islamic Azad University, Iran
Israa Shaker Tawfic,	Ministry of Science and Technology, Iraq
Iyad Alazzam,	Yarmouk University, Jordan
Jagadeesh HS,	APSCE (VTU), India
Jalal Srar,	Misurata University, Libya
Janaki Raman Palaniappan,	Brunswick Corporation, USA
Javier Gozalvez,	Universidad Miguel Hernandez de Elche, Spain
Jawad K. Ali,	University of Technology, Iraq
Jaymer Jayoma,	Caraga State University, Philippines

Jesuk Ko,	Universidad Mayor de San Andres, Bolivia
Jia Ying Ou,	York University, Canada
Jiong Li,	Space Engineering University, China
Jonah Lissner,	Technion - israel institute of technology, Israel
Jong-Ha Lee,	Keimyung University, South Korea
Jun Hu,	Harbin University of Science and Technology, China
Juntao Fei,	Hohai University, P. R. China
Kamel Benachenhou,	Blida University, Algeria
Kamel Hussein Rahouma,	Minia University, Egypt
Kamel Jemai,	University of Gabes, Tunisia
Kanstantsin MIATLIUK,	Bialystok University of Technology, Poland
Karim El Moutaouakil,	FPT/USMBA, Morocco
Karim Mansour,	University Salah Boubenider, Algeria
Karima Saidi,	Abbes Laghror University Khenchela, Algeria
Kazi Sultana Farhana Azam,	Friedrich-Schiller University Jena, Germany
Ke-Lin Du,	Concordia University, Canada
Keneilwe Zuva,	University of Botswana, Botswana
Khalid M.O Nahar,	Yarmouk University, Jordan
Khosrow Shafiei Motlagh,	Islamic Azad University, Iran
Kire Jakimoski,	FON University, Republic of Macedonia
Kiril Alexiev,	University of Diyala, Iraq
Kirtikumar Patel,	Hargrove Engineers and Constructors, USA
Klenilmar Lopes Dias,	Federal Institute of Amapa, Brazil
Koh You Beng,	University of Malaya, Malaysia
Kolla Bhanu Prakash,	KL University, India
Litao Guo,	Xiamen University of Technology, China
Ljiljana Trajkovic,	Simon Fraser University, Canada
Luca De Cicco,	Politecnico di Bari, Italy
Luis Gomez,	Universidad de Las Palmas de Gran Canaria, Spain
M.K.Marichelvam,	Mepco Schlenk Engineering College, India
Mabroukah Amarif,	Sebha University, Libya
Mahdi Abbasi,	Bu-Ali Sina University, Iran
Mahdi Sabri,	Islamic Azad University, Iran
Mahendra Kumar,	Sawangi (Megeh), India
Mahsa Mohaghegh,	Auckland University of Technology, New Zealand
Malka N. Halgamuge,	La Trobe University, Australia
Mamata Rath,	Birla Global University, India
Mamoun Alazab,	Charles Darwin University, Australia
Maniruzzaman,	Khulna University, Bangladesh
Manish Kumar Mishra,	University of Gondar, Ethiopia
Mario Versaci,	DICEAM - Univ. Mediterranea, Italy
Markos G. Tsipouras,	University of Western Macedonia, Greece
Mehdi Nezhadnaderi,	Islamic Azad university, Iran
Michail Kalogiannakis,	University of Crete, Greece
Mihai Carabas,	University POLITEHNICA of Bucharest, Romania
Mirsaeid Hosseini Shirvani,	Islamic Azad University, Iran
Mohamed A H Eleiwa,	EE Professor and Chairman, Egypt
Mohamed Hamlich,	Ensam UH2C, Morocco
Mohamed Ismail Roushdy,	Ain Shams University, Egypt
Mohammad A. Alodat,	Sur University College, Oman
Morteza Alinia Ahandani,	University of Tabriz, Iran
Mourad Chabane Oussalah,	University of Nantes, France

Shashikant Patil,	ViMEET Khalapur Raigad MS, India
Shicheng Zu,	Ericsson Panda Communication, China
Shing-Tai Pan,	National University of Kaohsiung, Taiwan
Shoeib Faraj,	Technical and Vocational University Of Urmia, Iran
Siarry Patrick,	Universite Paris-Est Creteil, France
Siddhartha Bhattacharyya,	Rajnagar Mahavidyalaya, India
Sidi Mohammed Meriah,	University of Tlemcen, Algeria
Sikandar Ali,	China University of Petroleum, China
Smain Femmam,	UHA University, France
Sofiane Bououden,	University Abbes Laghrour Khenchela, Algeria
Soha rawas,	Beirut Arab University, Lebanon
Souhila Silmi,	USTHB University, Algeria
Stefano Michieletto,	University of Padova, Italy
Subhendu Kumar Pani,	Krupajal Engineering College, India
Suhad Faisal Behadili,	University of Baghdad, France
Suhad Faisal,	University of Baghdad, Iraq
Sujatha,	School of Electronics Engineering, India
Sunanda Dixit,	BMS Institute of Technology and Management, India
Sun-yuan Hsieh,	National Cheng Kung University, Taiwan
Surajit Kundu,	NIT Sikkim, India
Taha Mohammed Hasan,	University of Diyala, Iraq
Taleb Zouggar Souad,	Oran 2 University, Algeria
Tamara Saad,	Baghdad College of economic sciences university, Iraq
Tanzila Saba,	Prince Sultan University, Saudi Arabia
Thaweesak Yingthawornsuk,	King Mongkut's University of Technology, Thailand
Titas De,	Data Scientist - Glance Inmobi, India
Toufik Bouden,	Mohammed Seddik Benyahia University of Jije, Algeria
Ugrasen Suman,	Devi Ahilya University, India
Umesh Kumar Singh,	Vikram University, India
Usman Naseem,	University of Sydney, Australia
V.Ilango,	CMR Institute of Technology, India
Venkata Duvvuri,	Oracle Corp & Purdue University, USA
W.R. Sam Emmanuel,	Nesamony Memorial Christian College, India
Wenwu Wang,	University of Surrey, UK
William R. Simpson,	Institute for Defense Analyses, USA
WU Yung Gi,	Chang Jung Christian University, Taiwan
Xiao-Zhi Gao,	University of Eastern Finland, Finland
Yasser M. Alginahi,	Department of Computer Science Adrian College, USA
Yew Kee Wong,	HuangHuai University, China
Yousef J. Al-Houmaily,	Institute of Public Administration, Saudi Arabia
Youssef Taher,	Center of Guidance and Planning, Morocco
Yuan Tian,	Nanjing Institute of Technology, China
Yuan-Kai Wang,	Fu Jen Catholic University, Taiwan
Yu-Chen Hu,	Providence University, Taiwan
Yuchen Zheng,	Shihezi University, China
Zakaria Kurdi,	University of Lynchburg, USA
Zhiwei Guo,	Chongqing Technology and Business University, China
Zhou RouGang,	HangZhou DianZi University, China
Zhu Shengxin,	Beijing Normal University, China
Zoltan Gal,	University of Debrecen, Hungary
Zoran Bojkovic,	University of Belgrade, Serbia

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Artificial Intelligence Community (AIC)

Soft Computing Community (SCC)

Digital Signal & Image Processing Community (DSIPC)

8th International Conference on Artificial Intelligence and Fuzzy Logic System (AIFZ 2022)

8th International Conference on Signal and Image Processing (SIGPRO 2022)

8th International Conference on Networks & Communications (NWCOM 2022)

3rd International Conference on Cloud Computing, Security and Blockchain (CLSB 2022)

3rd International Conference on Machine Learning Techniques and NLP (MLNLP 2022)

8th International Conference on Data Mining (DTMN 2022)

3rd International Conference on Big Data & IoT (BDIoT 2022)

8th International Conference on Computer Science, Engineering and Information Technology (CSITY 2022)

Uncertainty-Oriented Textual Marker Selection for Extracting Relevant Terms from Job Offers

Albeiro Espinal[12] and Yannis Haralambous[1] and Dominique Bedart[2] and John Puentes[1]

[1]IMT Atlantique, Lab-STICC, CNRS UMR 6285, Brest, France
[2]DSI Global Services, Le Plessis Robinson, France

Abstract

Automated resume ranking aims at selecting and sorting pertinent resumes, among those sent to answer a given job offer. Most of the screening and elimination process relies on the resumes' content, marginally including information of the job offer. In this sense, currently available resume ranking approaches lack of accuracy in detecting relevant information in job offers, which is imperative to assure that selected resumes are pertinent. To improve the extraction of relevant terms that represent significant information in job offers, we study the uncertainty-oriented selection of 16 textual markers – 10 obtained by examining the behaviour of expert recruiters and 6 from the literature – according to two approaches: fuzzy logistic regression and fuzzy decision trees. Results indicate that globally, fuzzy decision trees improve the F1 and recall metrics, by 27% and 53% respectively, compared to a state-of-the-art term extraction approach.

Keywords

Recruiter's Behavior Modeling, Relevant Term Extraction, Textual Relevance Marker Evaluation, Uncertainty Measure, Fuzzy Machine Learning.

1. Introduction

Job offers (JOs) and curriculum vitaes (CVs) are the documents through which recruiters and candidates interact, as part of a recruiting process. An important stage carried out by recruiters is the "Screening Phase" that evaluates the CVs of candidates to identify those who are qualified for a job. Analyzing both the main requirements of a new JO and the skills of the candidates expressed in their CVs can be very complex. This is specially the case when recruiters receive dozens or hundreds of candidates resumes [1]. In order to reduce such complexity, multiple artificial intelligence models have been developed to analyze and rank CVs for a given JO.

Although several models have been proposed, the automatic ranking of CVs remains a difficult task. In part, this is due to three issues that have rarely been examined in the literature. First, the most relevant information in the JO is not optimally identified, generating irrelevant rankings with respect to the essential requirements [2]. Secondly, under-representation of the changing organizational context surrounding JOs tend to break this type of systems over time [2]. Thirdly, since writing JOs engages human cognition, the expressed information is highly susceptible to uncertainty phenomena like ambiguity [3], which could render AI models ineffective [4]. Being still an active research field [5], the study of uncertainty and its characterization, is fundamental to investigate the extraction of relevant terms from JOs.

An organization's context to define a set of relevant textual markers based on recruiters' strategies to select significant JOs' information, and estimated the consistency of those markers was already studied [6]. Nevertheless a question remains concerning the quantitative evaluation

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 01-16, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121601

of identified markers' uncertainty, which is the goal of this work. Our study intends to assess the pertinence of automatically identified relevant JO terms, applying two machine learning models – fuzzy logistic regression and fuzzy decision trees – focused on the quantification of uncertainty. This article is organized as follows. Section 2 describes the related state of the art. We summarize some key aspects of our previous work in Sections 3 and 4. Section 5 describes the proposed uncertainty evaluation of textual markers. Experimental results are presented in Section 6. Discussion, conclusions and perspectives are presented in Sections 7 and 8.

2. State of the Art

CV ranking systems carry out three processing stages: CVs and JO pre-processing, representation, and automatic ranking of CVs in relation to the content of the JO. Those documents are pre-processed by extracting text from digital files (.pdf, .doc, .txt, among others). Then extracted texts can be standardized by eliminating noisy symbols, segmenting the documents, and making semantic annotations [1], as well as deleting stopwords [7]. Pre-processed documents can be represented based on n-gram models [1], bag-of-words [1], ontologies [8] and/or word embeddings [9]. From these representations, different approaches can be used to determine the most suitable CVs regarding a JO. They can rely on recruiters' feedback [1], neural architectures [9] and/or transformer models [10].

These methods, however, do not focus on extracting relevant information from the JO before ranking resumes. Some methods have proposed statistical and graph-based textual relevance markers for identifying significant terms in single documents [11] [12] [13].

Furthermore, uncertainty, a key concern of natural language processing, as automatic extraction of relevant information from individual documents [4], concerns the lack of information about an event. Three of the most studied approaches to determine uncertainty have been probability models [5], along with possibility theory and fuzzy logic models [14]. Contrary to probability-oriented models, fuzzy models assume that probability distributions cannot be obtained for fuzzy data. In this regard, linear and non-linear fuzzy machine learning models have been proposed to deal with uncertainty. Linear models as the fuzzy logistic regression are utilized to deal with uncertainty as fuzziness and not as randomness [4]. Also, non-linear models as fuzzy decision trees have been studied, including ambiguity and vagueness metrics to estimate uncertainty [3].

We propose to evaluate the uncertainty of textual markers that indicate the relevance of information in JOs based on recruiters' knowledge. The proposed evaluation compares fuzzy linear and non-linear machine learning methods, which are appropriate to investigate the uncertainty question, because of their possibilistic foundations at the crossroad of fuzzy sets and probability provide a simple and convenient setting for handling subjective tasks, as the automatic identification of the most relevant terms in JOs. Moreover, these types of models can be trained on small datasets to evaluate features relevance.

3. Representation of Job Offers

In order to evaluate the uncertainty of textual markers it is first necessary to specify the organizational context of JOs, analyze what is relevant for recruiters in this type of document, and extract textual markers that represent relevant information [6].

3.1. Organizational Context

The representation of societal contexts in machine learning models should be improved, allowing those models to become more adaptable to dynamic changes in organizations [15]. This is a critical aspect in our work, given that context influences strongly recruiters behaviors

[16]. We began thus by representing the recruiters' context before analyzing their strategies related to information relevance in JOs.

To this end, we used the UNC-method for representing organizational contexts, as specified in [17], by conducting an open dialogue with recruiters, specifying the entities and relationships that impact the JOs' life-cycle. As a result, the main entities, actors, processes, objectives, and organizational problems associated with JO management were identified. A pre-conceptual scheme was derived from this procedure and used for the construction of a mother-ontology, schematically described in the next section.

3.2. Ontology Derivation

We define a mother-ontology as a large ontology of module specifications. A mother-ontology was used to represent the main concepts and relationships inherent to the recruiters' context and JOs. Additionally, existent ontologies related to the particular organizational context were integrated into it. This was the case of the internal professional skills ontology of DSI Group which contains the specification of more than 36.000 professional skills, the european ontology of professional skills ESCO[1], the professional skills and job types frameworks of O*NET[2], CIGREF[3], and ROME[4], based on text-to-RDF-triple transformations [18]. The integration of these ontologies was achieved using a hybrid approach based on Bidirectional Encoder Representations from Transformers (BERT) [19], an analysis of terminological variation [20], and measures of ontology quality [23].

In this compound ontology, we specified also the structure of JOs in terms of concepts as sections, paragraphs, sentences, syntagms, terms, words, etc. Additionally, synonyms, meronyms and hyponyms were used to describe relationships between concepts. This enabled us to construct a more structured fuzzy model of the natural language contained in JOs by representing the basic constituents, as it has been suggested by [25]. An upper-view of the ontology is presented in Figure 1.

3.3. Analysis of Recruiters Viewpoints

Based on the organizational context representation using the previous ontology, we analyzed recruiters' strategies related to the selection of the most essential information in JOs. During the annotation process they highlighted the most relevant terms. To represent the description of each recruiter's observed actions, the controlled language proposed by [17] was used. It allows to represent actions sequentially, as triplets of the form <subject, verb, predicate>.

We categorized those actions as explicit (eg, <recruiter, selects, a term>) or implicit (eg, <recruiter, avoids, a term> or <recruiter, avoids, a JO section>). Once the annotations were described in a controlled manner, the Apriori algorithm was used to identify action sub-sequences that the recruiter performed systematically. These sub-sequences of actions describe behavioral patterns, formalized as semantic rules, using the mother-ontology described in section 3.2. Obtained rules represent textual relevance markers of information in JOs.

[1] https://esco.ec.europa.eu/en

[2] https://www.onetonline.org/

[3] https://www.cigref.fr/

[4] https://www.pole-emploi.fr/employeur/vos-recrutements/le-rome-et-les-fiches-metiers.html

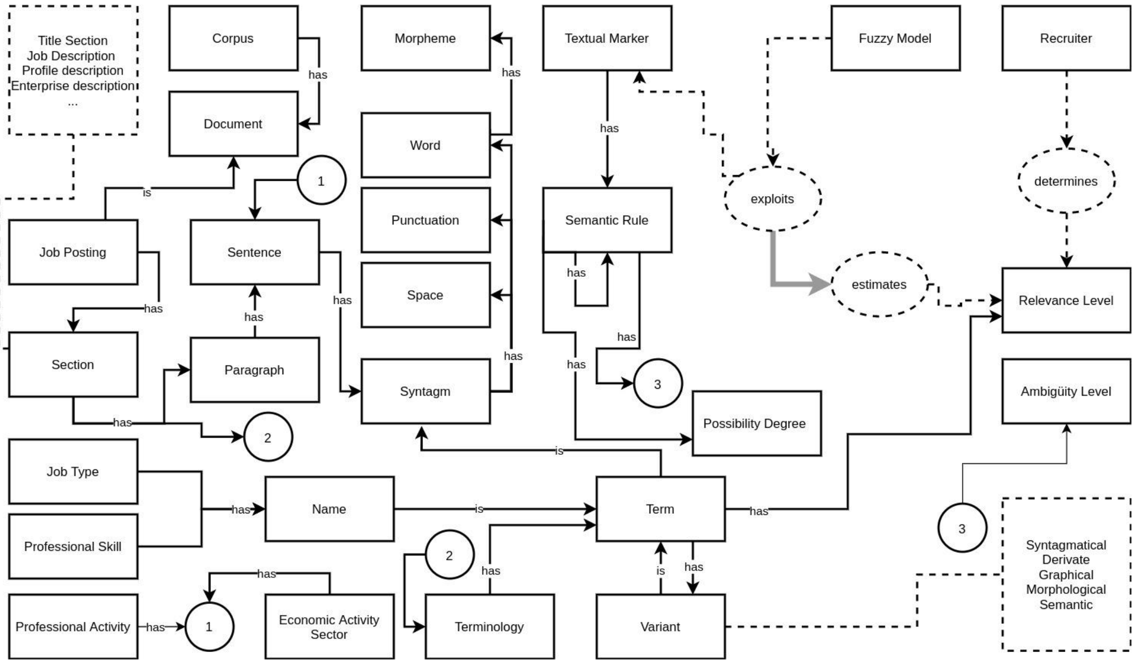


Figure 1. Upper-view of the mother-ontology created from the representation of the organisational context according to the principles of [17].

4. Textual Markers

In this section, we present briefly the evaluated textual markers and introduce the linguistic representation of JOs in our approach.

4.1. Initial Elements

Considering that terms are defined as functional classes of lexical units used in discourse [20], JOs' relevant terms were identified by the weirdness ratio that measures their termhood [20, 22]. Additionally, we extracted the JOs' terms by using the most frequent morphosyntactic rules of the language, identified on multiple corpora analysis [20], which are mostly nominal phrases. A JO is then represented by its terminology, and our approach aims to identify the most optimal set of textual markers.

Let d_i be a JO belonging to a corpus C and $T_{d_i} = \{t_1, t_2, ..., t_n\}$ the set of terms of d_i. Let $R_{d_i} \subseteq T_{d_i}$ be the *set of most relevant terms* in d_i. Each term $t_i \in R_{d_i}$ is considered as relevant under a possibility degree $\alpha_{t_{k,i}} \in [0, 1]$.

Let $A_{d_i} = \{a_1, a_2, ..., a_m\}$ be the set of sections of d_i (job description, profile details, etc). Each section a_i can be represented by a subset of terms from T_{d_i}. A term can belong to multiple sections. Let $E_{d_i} = \{e_1, e_2, ..., e_p\}$ be a set of qualifying adjectives and nouns that are linked to a subset of terms in T_{d_i} by syntax dependencies.

Let $O = \{o_1, o_2, ..., o_s\}$ be a set of ontologies (as the one presented in Section 3). Let $c_{o_s} = \{c_{s,1}, c_{s,2}, ..., c_{s,k}\}$ be the set of concepts of ontology o_s and $T_{c_j} = \{t_{j,1}, t_{j,2}, ..., t_{j,l}\}$ the set of terms lexically representing concept c_j in a given language.

4.2. Description of Textual Markers

In this section, we provide a summary of the derived textual markers [6] evaluated applying the proposed approach. Each marker provides a possibility degree for each JO's term of becoming relevant. Textual markers TM_1 to TM_{10} have been obtained from recruiters behaviors, while markers TM_{11} to TM_{16} correspond to those of the YAKE! (Yet Another Keyword Extraction) algorithm [12], found to be suitable, compared to other available algorithms in the literature. It is a domain-independent method applied in our case to identify potential relationships between textual markers and the context specificities of JOs.

4.2.1. Title Sections (TM_1)

"Any term in the title that resembles a term indicating professional skills or job types may potentially qualify as relevant."

Let $a_1 \in A_{d_i}$ be the title section of d_i. Let $t_{a1} = \{t_1, t_2, ..., t_u\}$ be the set of terms contained in a_1. Lexically, T_{c_j} is the set of terms that represent a professional skill or job type concept c_j in the ontology o_s. Therefore:

$$\forall t_k \exists c_j [c_j \in o_s \wedge t_k \in T_{c_j} \wedge t_k \in t_{a1}] \rightarrow t_k \in R_{d_i}, \tag{1}$$

with a possibility degree $\alpha_{t_{k,1}} \in [0, 1]$.

4.2.2. Terms Representing Professional Skills in a Job Description Section or Profile Description Section (TM_2)

Terms representing professional skills used in job descriptions or profile descriptions are more likely to be chosen as relevant terms. Let s_2 and s_3 be the sets of terms used in the job description section and the profile description section, respectively. Set $t_k \in T_{d_i}$. Let T_{c_j} be the set of terms used to represent a professional skill concept c_j in the ontology o_s. We request that:

$$\forall t_k \exists c_j ((t_k \in s_2 \vee t_k \in s_3) \wedge t_k \in T_{c_j}) \rightarrow t_k \in R_{d_i}, \tag{2}$$

with a possibility degree $\alpha_{t_{k,2}} \in [0, 1]$.

4.2.3. Relevance of Job Posting Sections (TM_3)

"As a general rule, recruiters are more likely to select terms from the title, job description, and profile description sections, rather than from other sections (company description, contract details, etc.)". As we don't require terms to be professional skills, this marker does not overlap with markers TM_1 and TM_2. Let $S = s_1 \cup s_2 \cup s_3 \subseteq T_{d_i}$, where: s_1 is the set of terms of the title section; s_2 is the set of terms of the job description section; and s_3 is the set of terms of the profile description section. Let $t_m \in T_{d_i} \cap S$. Then, we request that:

$$\forall t_m \forall t_n (t_m \in T_{d_i} \wedge t_n \notin S) \rightarrow (P(t_m \in R_{d_i}) > P(t_n \in R_{d_i})), \tag{3}$$

with a possibility degree $\alpha_{t_{k,3}} \in [0, 1]$. $P(t_* \in R_{d_i})$ represents the possibility of t_* being selected as a pertinent term.

4.2.4. Terms Dependent on Pertinence Expressions (TM_4)

“A relevant term is more likely to be one that bears a syntax dependency with a JO's syntagm.”

- Let $t_k \in T_{d_i} \cap T_{c_j}$ for some c_j.
- We define a “pertinent expression” e_m as a syntagm that the recruiter employed in the JO (i.e., *excellent* C# skills, *good understanding* of Kubernets). Assume that e_m is syntactically dependent with t_i. Specifically, let t_k be a qualifying adjective or a noun modifier directly dependent with e_m. Then:

$$\forall t_k \exists e_m (t_k \in T_{d_i} \wedge e_m \in E_{d_i} \wedge is\_dependent(t_k, e_m)) \rightarrow t_k \in R_{d_i}, \tag{4}$$

with a possibility degree $\alpha_{t_{k,4}} \in [0, 1]$.

4.2.5. Terms Used in Traces of Professional Activities Descriptions (TM_5)

“If a JO explicitly describes an interaction with a professional concept, a term representing that concept is more likely to be considered relevant.”

In a JO, a trace of a professional activity is a sentence that describes an action performed by a worker. Be $b_j \in d_i$ a trace of a professional activity description described by the set of terms T_{b_j}. We request that b_j contains at least one verb and one dependent object. As a result, the terms t_k that represent these objects will have a higher chance of being selected as relevant. Thus:

$$\forall t_k (t_k \in T_{b_j} \wedge is\_object(t_k, b_j)) \rightarrow t_k \in R_{d_i}, \tag{5}$$

with a possibility degree $\alpha_{t_{k,5}} \in [0, 1]$.

4.2.6. Terms Representing High Risk Professional Skills/Activities (TM_6)

In this marker, we aim to provide more relevance to terms that represent professional skills or activities on which an employee's mistake can adversely affect the company's economic performance. Value 0 indicates that a potential error will not significantly affect the economic activity, while value 1 indicates significant effects.

An ontology M describes the set of professional skills and activities of a given company. M contains a set of concepts $c_M = \{c_{M,1}, c_{M,2}, ..., c_{M,k}\}$. Recruiters manually assign a risk level $\epsilon_{c_{M,k}} \in [0, 1]$ to professional skills and activities.

Let s_j be a term in a JO d_i representing a professional skill or activity in M. As one of the concepts associated to s_j, let $c_{M,l}$ be the one with the highest risk level. When this risk level exceeds a threshold $\beta_{c_{M,l}}$, then s_j is selected as a pertinent term and:

$$\forall s_j \exists c_{M,l} (s_j \in T_{d_i} \wedge c_{M,l} \in M \wedge s_j \in T_{c_{M,l}} \wedge is\_greater\_than(\epsilon_{c_{M,l}}, \beta_{c_{M,l}}) \rightarrow s_j \in R_{d_i}, \tag{6}$$

with possibility degree $\alpha_{s_{j,6}} \in [0, 1]$.

4.2.7. Actions Expressed in Management JOs (TM_7)

The recruiter can identify what type of actions management JOs are required to perform. A management job might focus on team management, while another may involve accountability activities or even development tasks.

Be d_i a management JO. Based on 14,000 curriculum vitae, a Latent Dirichlet Allocation model was trained to detect management JOs. Let t_k be a verbal term of d_i. If t_k is part of the trace of a professional activity f_j and corresponds to the head of its syntactic tree, then this term may be relevant. We define it as follows:

$$\forall t_k \exists f_j (f_j \in d_i \wedge t_k \in f_j \wedge is\_management(d_i) \wedge \ is\_verb(t_k) \wedge is\_head\_of(t_k, f_j)) \rightarrow t_k \in R_{d_i}, \quad (7)$$

with a possibility degree $\alpha_{t_{k,7}} \in [0, 1]$.

4.2.8. BERT Semantic Similarity of Professional Skills (TM_8)

"If a *specific term* that represents a professional skill is semantically close (in the sense of BERT) to already discovered relevant terms, then it will be considered relevant."

Let $t_1 \in R_{d_i}$ and $t_2 \in T_{d_i}$. Let $f(t)$ be the specificity function of a term t defined as its relative frequency in a specific corpus C_s, divided by its frequency in a multi-language corpus C_L [20].

Furthermore, we define $g(t_1, t_2)$ as the BERT semantic similarity between two terms. Using a SBERT [24] model pre-trained on Wikipedia corpus, complex terms were semantically analyzed. As a result, this model was fine-tuned based on the following professional skill standards: CIGREF, e-CF, C2I, and ROME. We defined it as follows:

$$\forall t_1 \forall t_2 (t_1 \in R_{d_i} \wedge g(t_1, t_2) > 0) \rightarrow t_2 \in R_{d_i}, \quad (8)$$

with a possibility degree defined by the normalized equation:

$$\alpha_{t_2,8} = ||(1 - \alpha_{t_1}) * g(t_1, t_2) * f(t_2))||. \quad (9)$$

4.2.9. Relevance of the Economic Activity Sector (TM_9)

"Potentially relevant terms refer to the economic activities required by the job posting (e.g., finance, banks, aeronautics, etc.)". This implies that:

$$\forall t_k (t_k \in T_{d_i} \wedge is\_sector\_requirement(t_k)) \rightarrow t_k \in R_{d_i}, \quad (10)$$

with a possibility degree $\alpha_{t_{k,9}} \in [0, 1]$. In order to identify economic activity sectors, we aligned job posting terms and economic activity concept labels, provided by ESCO, O*NET, ROME, and CIGREF standards.

4.2.10. Professional Skill Prerequisites (TM_{10})

Assume there is a *prerequisite relation* between two professional skills c_1 and c_2 in an ontology o_i. Ontologies such as ESCO can be used to derive relations of this type. The possibility degree of c_1 will be inherited by c_2 if c_2 is a prerequisite of c_1 and c_1 is relevant (under a certain possibility degree).

$$\forall t_1 \forall t_2 \exists c_1 \exists c_2 (c_1 \in o_i \wedge c_2 \in o_i \wedge t_1 \in T_{c_1} \wedge t_2 \in T_{c_2} \wedge is\_prerequisite(c_1, c_2) \wedge t_1 \in R_{d_i}) \quad (11)$$

with a possibility degree $\alpha_{t_{k,10}} \in [0, 1]$ and $\alpha_{t_{k,10}}$ is equal to the possibility degree of $t_1 \in R_{d_i}$.

4.2.11. YAKE! Casing (TM_{11})

There is a tendency for upper-case terms to be more relevant. This YAKE! maker is defined as:

$$\forall t_k (t_k \in T_{d_i} \wedge is\_upper\_cased(t_k)) \rightarrow t_k \in R_{d_i} \quad (12)$$

The normalized YAKE! equation is used to calculate the possibility degree as:

$$\alpha_{t_{k,11}}(t_k) = \|\frac{max(TF(U(t_k)), TF(A(t_k))}{ln(TF(t_k))}\|, \quad (13)$$

where $TF(U(t_k))$ is the number of times that t_k appears uppercased, $TF(A(t_k))$ is the number of occurrences of t_k as an acronym (for details see) and $TF(t_k)$ is the term frequency.

4.2.12. YAKE! Term Position (TM_{12})

In this marker, the hypothesis is that terms that appear at the beginning of the document tend to be more pertinent.

$$\forall t_k (t_k \in T_{d_i} \wedge is\_position\_marker\_activated(t_k)) \rightarrow t_k \in R_{d_i}, \quad (14)$$

with a possibility degree obtained from the following normalized YAKE! equation:

$$\alpha_{t_{12}}(t_k) = \|ln(ln(3 + Median(Sent(t_k))))\|, \quad (15)$$

$Sent(t_k)$ is the set of positions of the sentences containing t_k.

4.2.13. YAKE! Term Frequency Normalization (TM_{13})

There is more relevance to the terms that are commonly used:

$$\forall t_k (t_k \in T_{d_i} \wedge is\_frequency\_marker\_activated(t_k)) \rightarrow t_k \in R_{d_i}, \quad (16)$$

The possibility degree is calculated based on the following normalized equation proposed by YAKE!:

$$\alpha_{t_{k,13}}(t_k) = \|\frac{TF(t_k)}{MeanTF + \sigma}\|, \quad (17)$$

where $TF(t_k)$ is the number of occurrences of t_k, which is balanced by the mean and standard deviation of frequency.

4.2.14. YAKE! Term Relatedness to Context (TM_{14})

This YAKE! marker is based on the following hypothesis: “The more terms co-occur on both sides of a candidate term t, the less significant that term is”:

$$\forall t_k (t_k \in T_{d_i} \wedge is\_relatednes\_activated(t_k)) \rightarrow t_k \in R_{d_i}, \quad (18)$$

with a possibility degree obtained from the normalized YAKE! equation:

$$\alpha_{t_{k,14}} = \|1 + (DL + DR \cdots) * \frac{TF(t_k)}{maxTF}\|, \quad (19)$$

where

$$DL[DR] = \frac{|A_{t,w}|}{\sum_{k \in A_{t,w}} CoOccur_{t,k}} \tag{20}$$

In a window of size w, $|A_{t,w}|$ corresponds to the number of different terms, and TF is the term frequency.

4.2.15 YAKE! Different Sentences (TM_{15})

“A term's relevance depends on how frequently it appears within different sentences”, defined as:

$$\forall t_k (t_k \in T_{d_i} \wedge is\_sentences\_marker\_activated(t_k)) \rightarrow t_k \in R_{d_i}, \tag{21}$$

with a possibility degree obtained from the normalized equation:

$$\alpha_{t_{k,15}} = \| \frac{SF(t_k)}{\#Sentences} \|, \tag{22}$$

where $SF(t_k)$ is the number of sentences containing t_k and $\#Sentences$ is the total number of sentences of d_i.

4.2.16. YAKE! Overall Score (TM_{16})

Based on markers $TM_{11}, TM_{12}, TM_{13}, TM_{14}\, and\, TM_{15}$ proposed by YAKE!, we include its global relevance score. Let $t_k \in d_i$. A term is considered as “possibly relevant” if it’s predicted as such by the overall score:

$$\forall t_k (t_k \in T_{d_i} \wedge is\_predicted\_by\_yake(t_k)) \rightarrow t_k \in R_{d_i}, \tag{23}$$

with a possibility degree $\alpha_{t_{k,16}} \in [0, 1]$.

5. Evaluation of Textual Markers

Two factors should be considered regarding the recruiters’ annotations of job offers. Firstly, it is a classification task, since it consists on determining whether or not each term of a JO is relevant to describe its essential content. Being a classification task, it can be understood as a rational action that an expert recruiter takes according to his/her knowledge [3]. Secondly, the act of annotating documents can be thought of as an inference process that recruiters undertake when reading the JO. Therefore, their annotations may be highly subject to cognitive uncertainties, which should be integrated to natural language processing tasks [4]. In the following two sections, we present the two uncertainty-oriented models, applied to the evaluation of textual markers derived from recruiters’ strategies.

5.1. Preliminary Definitions

Let $U = t_1, t_2, ..., t_m$ be the set of terms of a JO, where m represents the number of terms extracted. Each JO term t_m can be described by a set of relevance textual markers (TM_k) derived from recruiters strategies and existent literature. We denote them as $I(k) = \{TM_1, TM_2, ..., TM_k\}$. Therefore, each term t_m can be represented in the following form:

$$(x_{i0}, x_{i1}, ..., x_{ij}, \widetilde{Y}_i),\ 1 \leq i \leq m \tag{24}$$

where x_{ij} corresponds to a possibility degree obtained from textual marker j for the term i of being a relevant term. $\tilde{Y}_i$ represents the recruiter’s annotation on this term which is inherently influenced by uncertainties (as such, we consider it an estimation $\tilde{Y}_i$ of the actual truth Y_i).

On the other hand, we define the fuzzy set C that aims to model the relevance levels of the terms that the recruiters identify in the JOs. C is composed of a membership function μ_C that allows to fuzzify the annotations made by the recruiters on the JOs. Furthermore, we define that the set C is composed of two fuzzy subsets: C_1 which represents the relevance levels of the relevant terms and C_2 which represents the relevance levels of the non-relevant terms. These functions have been modeled using triangular functions whose support covers the range (0,1). In addition, we define the fuzzy set R (resp. R_1, R_2), contained in C (resp. C_1, C_2), and obtained after fuzzifying the annotations made by the recruiters. In the following sections, we present how the linear – fuzzy logic logistic regression – and non-linear – fuzzy decision tree –, approaches were applied to assess the uncertainty of relevant textual markers.

5.2. Linear Evaluation: Fuzzy Logistic Regression

Be $t = \{t_1, t_2, t_3 \ldots t_m\}$ the set of terms of the JO. We assume that these terms can be represented as a linear combination of the set of textual markers $I(k)$. Applying the fuzzy logistic regression algorithm [14], let $\mu_i \in \{C_1(pertinent\ term), C_2(non\ pertinent\ term)\}$ be the recruiter’s annotation on the ith term of a job posting. We estimate the parameter $\tilde{u}_i$ from the ratio $\frac{\tilde{\mu}_i}{1-\tilde{\mu}_i}$. In our context, $\frac{\tilde{\mu}_i}{1-\tilde{\mu}_i}$ can be interpreted as the possibility of a term of not being relevant in relation to the possibility of being relevant, or vice versa. Therefore, the model is [14]:

$$\tilde{W}_i = ln\frac{\tilde{u}_i}{1-\tilde{u}_i} = A_0 + A_1 x_{i1} + \ldots + A_n x_{in}, i = 1, \ldots, m \qquad (25)$$

where $\tilde{W}_i$ is the estimated output that can be transformed back to $\tilde{u}_i$ by the extension principle and A_i=(a_i, s_i) represents a triangular fuzzy and symmetrical number with center a_i and spread s_i.

5.3. Non Linear Evaluation : Fuzzy Decision Trees

In order to train the fuzzy decision tree, we fuzzify each textual marker by applying a membership function μ_{TM_k}, built equivalently to μ_C, but taking into account the specific codomain of each marker TM_k. We define that this fuzzification represents an evidence E_k. From the fuzzification of each textual marker and recruiters’ annotations, we estimate the possibility of representing the fuzzified recruiters’ annotations R in light of the evidence E_k. In particular, we evaluate how ambiguous the following implication is: If E_k Then R. Multiple measures can be used to evaluate this implication [3]. We applied the subsethood measure to estimate how much the evidence E_k implies the experts’ classification R, according to:

$$S(E_k, R_i) = \frac{M(E_k, R_i)}{M(E_k)} = \frac{\sum_{t \in U} min(\mu_{E_k}(t), \mu_{R_i}(t))}{\sum_{t \in U} \mu_{E_k}(t)} \qquad (26)$$

In relation to recruiters’ strategies and viewpoints, we determine whether a term is relevant R_1 or not R_2 making use of:

$$\pi(R_i|E_k) = \frac{S(E_k,R_i)}{max(S(E_k,R_1),S(E_k,R_2))} \tag{27}$$

As possibility is intrinsically related to the concept of ambiguity [3], there is less ambiguity when we can clearly determine whether a term is relevant or not. From $\pi(R|E_k)$, we estimate the ambiguity level associated to marker TM_k linked to the evidence E_k as:

$$G(E_k) = g(\pi(R|E_k)) = \sum_{i=1}^{n} (\pi_i^* - \pi_{i+1}^*) ln(i) \tag{28}$$

where $\pi^* = \{\pi_1^*, \pi_2^*, ..., \pi_n^*\}$ is the possibility distribution $\pi(R|E_k)$ permuted and sorted so that $\pi_i^* \geq \pi_{i+1}^*$ for $i \in \{1, ..., n\}$ and $\pi_{n+1}^* = 0$.

Due to the fact that we evaluate ambiguity by considering whether a term is relevant (R_1) or not (R_2) based on TM_k, $n = 2$. Subject to this ambiguity function, we can estimate the extent to which it can be clearly inferred that a term is pertinent or not, according to I_k. Therefore, $ln(n)$ indicates maximum ambiguity and 0 represents no ambiguity [3]. To train the fuzzy three, our final step is to replace the classical information entropy measure with the previously presented ambiguity metric. In the case of complex evidences E_k composed by subsets of evidences, the ambiguity is estimated using the partitioning approach [3].

6. Experimental Results

A test of our approach was conducted at DSI Group's recruitment department. In total, 5 recruiters participated in our experiment and we refer to them as A, B, C, D, and E. These recruiters had in-depth knowledge of the essential JOs' requirements they manipulated within the setting of this experimentation.

6.1. Procedure

As indicated in section 3.1, our experimentation began with the representation of the organizational context surrounding JOs, based on interviews with recruiters. From this procedure, we derived the ontology illustrated in section 3.2. Then, we asked recruiter A, the director of the human resources department, to describe the most relevant requirements of five JO under his responsibility. In recruiting a candidate, relevant requirements are those that do not allow for any flexibility.

Using expert A's strategies for selecting the most essential information in each job opening we derived relevance textual markers from his strategies. Generally, the annotated terms relate to professional skills, and to a lesser extent, location and availability, among others. Once the textual markers were derived conforming to recruiter A findings, we invited the other four recruiters (B, C, D and E), to determine whether the strategy derived from recruiter A's behavior was valid or not, to analyze other CVs. This evaluation process was executed as follows:

- Recruiters B, C, D, and E annotated JOs that they had managed. We obtained a total of 25 annotated documents. On average, each job posting contained 100 terms of interest, out of which between 4 to 10 terms were annotated as relevant. A first dataset of 2501 terms was generated.
- To train the fuzzy models, a second dataset was generated using the random undersampling RUSBoost algorithm [22]. A dataset of 500 terms, with 35% relevant and 65% non relevant terms was obtained.
- Both the linear and non-linear fuzzy models were trained on 70% of the second dataset and tested on the remaining 30%. We used stratified sampling to guarantee the

proportion of relevant and non relevant terms on each dataset. Additionally, we examined the reliability of the resulting models by using a stratified 10-fold cross-validation.

- Both fuzzy models were compared to a state-of-the-art term extraction approach. For each annotated JO, we assessed the suitability of each model, based on the precision@K, recall@K, and F1-score@K metrics (where N represents the number of terms annotated by the recruiter).
- Model evaluations were done with the remaining terms of the first dataset, after the terms of the second dataset used for training were excluded. The training procedure allowed to obtain the best model avoiding overfitting and guaranteeing a maximal variance of the training samples. Finally, the evaluation procedure for measuring the precision@K, recall@K, F1-Score@k metrics had as a goal to confront the trained models to a much more realistic setting with a significant amount of non relevant terms.

6.2. Example of an Annotated Job Offer

Below, we present a summary view of an example JO annotated (with relevant terms in bold) by recruiter B.

***BI / BO** Analyst M/W*
Company Description...(it contains 121 words)
Job description... (it contains 89 words)
Profile Description... (it contains 69 words)
You hold a Computer Engineering degree. You have technical skills such as:
- Business Objects platform
*- **Mastery of the SQL language**, and the use of databases (**SAP IQ / IBM DB2**)*
Knowledge of Stambia ETL or Oracle. Data Integration would be appreciated
Good interpersonal skills, dynamism, spirit of synthesis, proactive,
and team spirit are qualities that characterize you.
Job experience: Minimum 2 years. Position location: Metz-57. Geolocatable: Yes.

Table 1. *Top $N = 5$ terms predicted by the Fuzzy Logistic Regression and Decision Tree.*

	Fuzzy Logistic Regression			**Fuzzy Decision Tree**		
#	**Term**	**Score**	**Interval**	**Term**	**Ambiguity %**	**Relevance Score**
1	DSI	0.98	$\pm\ 0.02$	BI	9	0.97
2	Mastery the SQL Language	0.93	$\pm\ 0.09$	BO	9	0.97
3	Enterprise Activity	0.91	$\pm\ 0.15$	Mastery of the SQL Language	16	0.87
4	BI	0.87	$\pm\ 0.16$	SAP IQ	28	0.71
5	SAP IQ	0.87	$\pm\ 0.16$	Technical Skill	25	0.69

Table 1 presents the top N=5 terms predicted by the fuzzy logistic regression and decision tree models on the example JO, as well as the relevance scores of each term, with the associated intervals and ambiguity levels. Some predicted terms (like DSI and Enterprise Activity) are part of the company/job description sections. In this case, both syntactically and semantically, the decision tree model predicts closely terms that are annotated by recruiters.

6.3. Experimentation

Table 2 presents the results of our experiments. All tests were done applying the fuzzy logistic regression (FLR) and fuzzy decision tree (FDT) approaches. We trained each model using state-of-the-art textual markers [E], the proposed context-driven textual markers [R], and combining the two textual markers extraction procedures [R+E]. As indicated by the metrics, the fuzzy decision tree results are significantly better than the fuzzy logistic regression and the YAKE! algorithm. We also evaluated the algorithms proposed by [11] [13], which under-performed YAKE!. The fuzzy decision tree improved the best results of the state-of-the-art approach from 27% to 53%, being 78% for Recall@2N the highest performance. Note that the state-of-the-art textual markers were adapted to the specific context of JOs through the training process.

Table 2. Precision, recall, and F1-score results of each method tested on 25 JOs (FLR: fuzzy logistic regression; FDT: fuzzy decision tree; [E]: state-of-the-art textual markers; [R]: proposed context-driven textual markers; [R+E]: combination of state-of-the-art and proposed context-driven textual markers.

Metric/Model	**YAKE!**	**FLR[E]**	**FDT[E]**	**FLR[R]**	**FDT[R]**	**FLR[R+E]**	**FDT[R+E]**
Precision@N, Recall@N and F1-Score@N[5]	0.10	0.16	0.19	0.24	0.38	0.41	**0.53**
Recall@2N	0.25	0.33	0.40	0.42	0.57	0.62	**0.78**
Precision@2N	0.12	0.16	0.20	0.21	0.28	0.31	**0.39**
F1-Score@2N	0.16	0.22	0.27	0.28	0.37	0.41	**0.52**

Table 3 presents the coefficient values for each of the textual markers, based on the obtained models. A classical logistic regression was also trained, to include a complementary well-known model. Evaluation of the textual markers' ambiguity applying the fuzzy decision tree reveals interesting aspects of how relevant terms are identified. For instance, low ambiguity appears for indicators TM_1, TM_{12}, and TM_{16}, indicating that: recruiters tend to take into account relevant terms in job titles (according to TM_1); terms appearing at the beginning of the document tend to be relatively relevant (in agreement with TM_{12}'s), which could be due to the company description section appearing at the beginning in some JOs; because of YAKE! features, often highly irrelevant terms are predicted as relevant (as reported by TM_{16}), being an estimation of counter-relevance of terms in our context.

[5] Recall@N, Precision@N and F1-Score@N are equivalent at N.

Table 3. Individual uncertainty evaluation of the 16 extracted textual markers applying classic logistic regression (CLR), fuzzy logistic regression (FLR), and fuzzy decision tree (FDT). Coef.: CLR coefficients, SE: CLR standard errors, Coef. A: center of the triangular fuzzy number, Coef. S: spread of the triangular fuzzy number.

	CLR			**FLR**		**FDT**
Textual Marker	Coef.	SE	p-value	Coef. A	Coef. S	Ambiguity %
TM_1	1.18	0.67	0.078	0.33	<0.001	12
TM_2	4.02	0.52	< 0.001	3.40	<0.001	40
TM_3	2.66	0.81	< 0.001	1.23	<0.001	26
TM_4	1.66	0.52	0.002	1.00	<0.001	17
TM_5	2.30	0.56	< 0.001	1.61	<0.001	18
TM_6	1.48	0.65	0.023	0.03	<0.001	9
TM_7	-0.41	0.63	0.512	0.63	<0.001	8
TM_8	1.81	0.53	< 0.001	1.08	<0.001	13
TM_9	-0.30	0.66	0.647	0.71	<0.001	8
TM_{10}	1.02	0.68	0.132	0.26	<0.001	8
TM_{11}	1.09	0.45	0.015	0.81	<0.001	39
TM_{12}	-0.56	0.26	0.029	-0.85	<0.001	19
TM_{13}	-0.27	0.63	-0.436	0.68	<0.001	31
TM_{14}	0.12	0.10	0.246	-0.02	<0.001	20
TM_{15}	3.87	2.73	0.160	1.71	<0.001	35
TM_{16}	1.86	0.91	0.041	0.41	<0.001	5
Intercept	-4.51	0.86	< 0.001	-2.48	0.730	

7. Discussion

Uncertainty evaluation is crucial to improve the identification of relevant terms extracted automatically from JOs. Our work proposes an analysis of possibility and uncertainty metrics, to assess the relevance of identified textual markers.

The classical logistic regression has a R^2 value of 0.64, which indicates a relative strong fit. This value was used as a convenient but not decisive indicator (because of the data uncertainty), revealing to which degree the introduction of the context-driven markers helped to better describe the recruiters viewpoints about what is relevant in JOs, from a statistical point of view. Moreover, our hypothesis that a probabilistic model of the recruiters' annotations was not sufficiently appropriate, is likely to be confirmed by the p-values of the classic logistic regression. According to the coefficients of the fuzzy logistic regression, recruiter-oriented indicators, TM_2, TM_3, TM_4, TM_5, and TM_8 seem to be the most pertinent contextual markers.

We noticed that marker TM_8 (similarity of terms with important skills) induces relevant terms corresponding to false-positives, strongly related to the JO's context (e.g. the term "Technical

Skill" predicted in section 6.2). Regarding the intercept value of the FLR by applying the extension principle [14], the *possibility* of predicting a term as highly relevant is centered on 8% if all its textual markers values are zero, which is a more pertinent assumption due the uncertainty of recruiters viewpoints. Instead the intercept of CLR model gives a *probability* centered on 1%, indicating that even if all the regressor variables are zero, there is a level of uncertainty still not described, associated to the recruiters viewpoints of information relevance.

The applied fuzzy models appear to be better suited to handle considerable uncertain information [4] communicated by recruiters. According to obtained results, the fuzzy decision tree shows a better performance, implying its feasible alignment with recruiters' strategies. This is supported by the fact that the fuzzy decision tree obtained a better F1-Score using only the context-driven markers, the context-independent markers, and both types of markers combined. Specifically, we observed that multiple decision rules obtained after training the fuzzy decision tree match previously behaviors observed in recruiters. The following rule is an example: "If it is highly possible that a term in the title represents a professional skill or job type (TM_1) and if it is highly possible that it represents a professional skill mentioned in the job or profile description sections (TM_2), then it is highly possible that such term is relevant."

We also observed that some domain-independent markers are correlated to the context of JOs. For instance, the TM_{11} marker is associated with the behavior of recruiters who capitalize terms representing professional skills, which are generally relevant to JOs. Despite its importance, such a marker could also be ambiguous (39%), which is consistent because capitalization does not necessarily imply importance. Globally, our results indicate that the most pertinent textual markers are TM_2, TM_3, TM_4, TM_5, TM_8, TM_{11} and TM_{12}.

8. Conclusions and Perspectives

In this study, we evaluated two fuzzy models – linear and non-linear – for assessing the uncertainty of textual markers, in terms of ambiguity, with respect to recruiters' knowledge. Those textual markers serve to extract automatically relevant terms that are appropriate to model the information in JOs. It is therefore likely that reliable textual markers can be identified according to ambiguity. Possibility intervals and ambiguity scores provide flexibility to the evaluation process centered on uncertain information provided by experts, within a specific organizational context, with the potential of being adapted to other JOs' organizational contexts. In general, textual markers derived from recruiters' strategies were more pertinent than those extracted from the literature, although results improved significantly when both were combined.

These results provide further support to the suggestion that machine learning systems should systematically include an organizational context layer representation, which in our case certainly improved the evaluation of textual markers. The scope of this study was mainly limited in terms of the corpus size and the modeled aspects of the organizational context. Further research is therefore still required. It will be necessary to examine a larger corpus in order to determine whether the selected textual markers can be applied to different organizational contexts. Additionally, a question remains about the suitability of uncertainty measures to particularities of different organizations and the impact of organizational changes in the evaluation of textual relevance markers.

References

[1] L. A. Cabrera-Diego, M. El-Béze, J. M. Torres-Moreno, B. Durette, 'Ranking résumés automatically using only résumés: A method free of job offers', *Expert Systems with Applications* 123, 91–107, 2019.

[2] J. Martinez-Gil, A. L. Paoletti, M. Pichler, 'A Novel Approach for Learning How to Automatically Match Job Offers and Candidate Profiles', *Information Systems Frontiers* 22(6), 1265–1274, 2020.

[3] Y. Yuan, M. J. Shaw, ‘Induction of fuzzy decision trees’, *Fuzzy Sets and Systems*, 69(2), 125–139, 1995.

[4] E. Pavlick και T. Kwiatkowski, ‘Inherent Disagreements in Human Textual Inferences’, *Transactions of the Association for Computational Linguistics* 7, 677–694, 2019.

[5] M. Abdar, F. *Pourpanah, S. Hussain, D. Rezazadegan, L. Liu*, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U.R. Acharya, V. Makarenkov, S. Nahavandi, ‘A review of uncertainty quantification in deep learning: Techniques, applications and challenges’, *Information Fusion 76*, 243–297, 2021.

[6] A. Espinal, Y. Haralambous, D. Bedart, J. Puentes, ‘An Ontology-Based Possibilistic Framework for Extracting Relevant Terms from Job Advertisements’, In *2022 International Conference on Fuzzy Computation Theory and Applications (FCTA), Accepted for Publication*, 2022.

[7] P. K. Roy, S. S. Chowdhary, R. Bhatia, ‘A Machine Learning approach for automation of Resume Recommendation System’, *Procedia Computer Science* 167, 2318–2327, 2020.

[8] D. Çelik, ‘Towards a semantic-based information extraction system for matching résumés to job openings’, *Turkish Journal of Electrical Engineering and Computer Sciences* 24(1), 141–159, 2016.

[9] C. Zhu, H. Zhu, F. Xie, P. Ding, H. Xiong, C. Ma, P. Li, ‘Person-Job Fit: Adapting the Right Talent for the Right Job with Joint Representation Learning’, *ACM Transactions on Management Information Systems* 9, 1–17, 2018.

[10] X. Wang, Z. Jiang, L. Peng, ‘A Deep-Learning-Inspired Person-Job Matching Model Based on Sentence Vectors and Subject-Term Graphs’, *Complexity* 2021, 1–11, 2021.

[11] A. Zehtab-Salmasi, M.-R. Feizi-Derakhshi, και M.-A. Balafar, ‘FRAKE: Fusional Real-time Automatic Keyword Extraction’. 2021.

[12] R. Campos, V. Mangaravite, A. Pasquali, A. M. Jorge, C. Nunes, και A. Jatowt, ‘YAKE! Collection-Independent Automatic Keyword Extractor’, In *Advances in Information Retrieval*, 806–810, 2018.

[13] R. Dagli, A. M. Shaikh, H. Mahdi, και S. Nanivadekar, ‘Job Descriptions Keyword Extraction using Attention based Deep Learning Models with BERT’, In *3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. 1–6, 2021.

[14] S. Pourahmad, S. M. T. Ayatollahi, S. M. Taheri, Z. H. Agahi, ‘Fuzzy logistic regression based on the least squares approach with application in clinical studies’, *Computers and Mathematics with Applications*, 62(9), 3353–3365, 2011.

[15] D. Martin Jr, V. Prabhakaran, J. Kuhlberg, A. Smart, W. S. Isaac, ‘Extending the Machine Learning Abstraction Boundary: A Complex Systems Approach to Incorporate Societal Context’. 2020. arXiv 2006.09663.

[16] J. A. Breaugh, ‘Employee Recruitment’, *Annual Review of Psychology* 64, 389–416, 2013.

[17] C. M. Zapata Jaramillo, F. Arango Isaza, ‘The UNC-method: a problem-based software development method’, *Ingeniería e Investigación* 29, 69–75, 2009.

[18] M. Somodevilla García, D. Vilariño Ayala, I. Pineda, M. Somodevilla García, D. Vilariño Ayala, I. Pineda, ‘An Overview of Ontology Learning Tasks’, *Computación y Sistemas* 22(1), 137–146, 2018.

[19] S. Neutel, M. de Boer, ‘Towards Automatic Ontology Alignment using BERT’, In *AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering*, 2021.

[20] D. Cram, B. Daille, ‘Terminology extraction with term variant detection’, In *Proceedings of ACL-2016 system demonstrations*, 13–18, 2016.

[21] K. T. Frantzi, S. Ananiadou, J. Tsujii, ‘The C-value/NC-value Method of Automatic Recognition for Multi-word Terms’, Research and Advanced Technology for Digital Libraries 1513, 585–604, 2002.

[22] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, A. Napolitano, ‘RUSBoost: A Hybrid Approach to Alleviating Class Imbalance’, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 40(1), 185–197, 2010.

[23] S. Mc Gurk, C. Abela, J. Debattista, ‘Towards Ontology Quality Assessment’. 2017. Http://ceur-ws.org/Vol-1824/ldq paper 2.pdf.

[24] N. Reimers, I. Gurevych, ‘Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks’, In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019, 3982–3992.

[25] V. Novák, ‘Fuzzy logic in natural language processing’, *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2017, 1–6.

YOLOV5, YOLO-X, YOLO-R, YOLOV7 PERFORMANCE COMPARISON: A SURVEY

Ismat Saira Gillani[1], Muhammad Rizwan Munawar[2],
Muhammad Talha[3], Salman Azhar[4], Yousra Mashkoor[5],
Muhammad Sami uddin[6] and Usama Zafar[7]

[1]Department of Computer Science, Columbus State University, USA
[2]Department of Computer Science, COMSAT University Pakistan
[3]Department of Electrical Engineering, GC University Faisalabad, Pakistan
[4]Department of Creative Technology, Air University Pakistan
[5]Department of Computer Science, NED University Pakistan
[6]National University of Technology (NUTECH) Islamabad, Pakistan
[7]Department of Computer Science, GC University Faisalabad, Pakistan

ABSTRACT

YOLOv7 algorithm have taken the object detection domain by the storm as its real-time object detection capabilities out ran all other previous algorithms both in accuracy and speed [1]. *YOLOv7 advances the state of the art results in object detection by inferring more quickly and accurately than its contemporaries. In this paper, we are going to present our work of implementing this SOTA deep learning model on a soccer game play video to detect the players and football. As the result, it detected the players, football and their movement in real time. We also analyzed and compared the YOLOv7 results against its previous versions including YOLOv4, YOLOv5 and YOLO-R. The code is available at: https://github.com/RizwanMunawar/YOLO-RX57-FPS-Comparision*

KEYWORDS:

Deep Learning, Object detection, Image segmentation, Instance segmentation, Network Architecture.

1. INTRODUCTION

Real-time prediction of the presence of one or more objects, along with their classes and bounding boxes, is the task of computer vision that has taken the industry by storm. Object detection can use a neural network to classify and localize an object in the image. Benefitting from this capability, there is a tremendous amount of work that is being done in the different streams of life from facial recognition to autonomous driving cars, security applications and robotics [2]. Modern detectors have been in the development to identify the objects in higher frame rate. Recently, it has been seen how vehicles candrive on their own once put on the auto-pilot mode. However, their applications are not limited to these fields only. In this paper, state of the art YOLOv7 model will be utilized for real-time detection of the players and football movement. The results of this implementation were quite fascinating as there was highest accuracy seen along with speed. Generally, object detection methods are of two types; **a.** single-stage detectors, **b.** multi-stage detectors (Figure 1). The former methods directly deal with the execution speed whereas the latter focus more on the accuracy of the model and are evaluated in

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 17-28, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121602

MAP metric. Before YOLO, two stage detectors *(R-CNN, Fast R-CNN, and Faster R-CNN)* demonstrated the SOTA results in terms of accuracy [3].

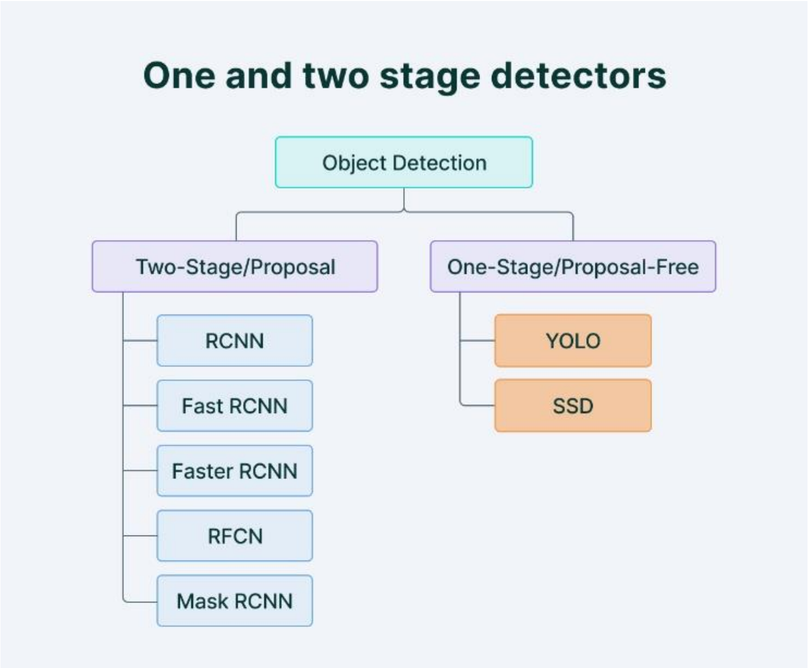


Figure 1. Comparison between One-Stage and Two-Stage Detectors

However, there is always a trade-off between speed and accuracy among these methods. When applied to real-time data, RCNN showed better accuracy as compared to RFCN which yielded more speed. Later on, YOLO model replaced other SOTA algorithms because of its speed and accuracy. YOLOv2 achieved results with a reasonable 78.6% *MAP* on *VOC 2007+2012* at 40 FPS [4]. Nevertheless, over the years many YOLO versions were developed and in each version there was a speed accuracy trade off.

Moving forward, the YOLOv5, YOLO-R and YOLOv7 will be talked about respectively. Firstly, the limitations of the first two models will be outlined and then improvements and structure of YOLOv7 will be discussed. Afterwards, we will compare the performance of all three models to analyze which one is the most accurate.

2. Related Work

The YOLO algorithm uses convolutional neural networks (CNN) to quickly identify objects. The approach just needs one forward propagation through a neural network, as the name would imply, to detect objects.

2.1. YOLOv5

For feature extraction from photos made up, it uses CSPDarknet as the foundation. In order to aggregate the features and pass them on to Head for prediction, it creates a feature pyramids network using PANet. Intuitively, the data is fed to the Backbone, which is nothing but aCSPDarknet, where feature extraction is happened and then they are passed to the Neck, aPANet, where the feature fusion takes place and then lastly, YOLO layer predicts the detection output in terms of class, location, confidence score and size [5]. *(Figure 3)*

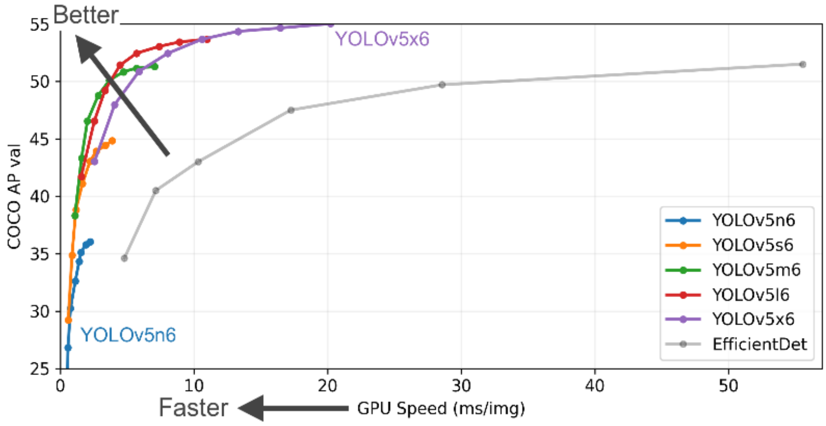


Figure 2. Comparison of Different YOLOv5 Variants

During training of this model, Leaky ReLU, sigmoid activation, SGD, and ADAM are available as optimizer choices in YOLOv5. It makes use of Logits loss and Binary cross-entropy. Just like every other model, it also has multiple variants and these varieties have a size and inference time trade off. The lightest variant YOLOv5s takes up to only 14MB however it lacks seriously in accuracy as compared to the YOLOv5x that has size of 168MB but is the most accurate in this family. It can be seen that YOLOv5x shows *48% AP* at 5ms per image as compared to YOLO5s that demonstrates *45% AP* at 5ms per image (*Figure 2)*.

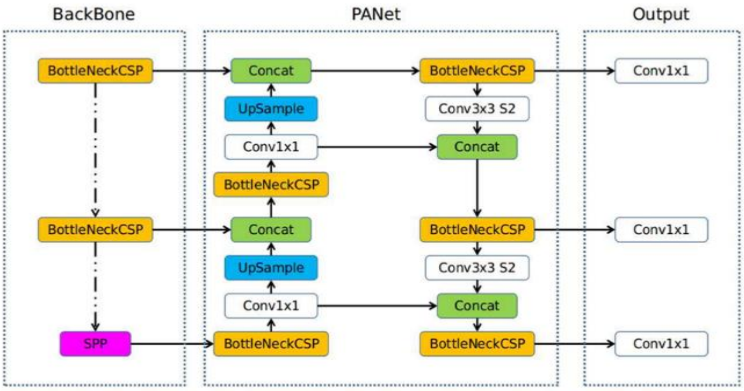


Figure 3. Network Architecture of YOLOv5 *[5]*

2.2. YOLO-R

This variant had the best inference time and accuracy among all YOLO models including YOLOv5. It proposes a single neural network that does multiple tasks like prediction, multi-tasking learning and feature calibration. YOLOR makes the most of Explicit and Implicit knowledge to build up the model that performs multi-label classification, detection, feature embedding all at once. This model learns from both the given data and the input (explicit knowledge) and the data learnt from the past experiences (implicit knowledge) [6].

A total of three processes that include kernel space alignment, prediction refinement and Convolutional Neural Network (CNN) make this architecture functional. CNNs try to retrieve an output according to an input. However, YOLOR yields both the CNNs that learn how to extract the output and all the possibilities of different outputs that could be, instead of just one result *(Figure 4)*.

When compared with YOLOv4, YOLOR-E6 delivers the better AP value with *5%* less computation. It use *10%* less parameters and improved inference speed by *15%*. In case of U5R5-X6, YOLOR demonstrates best AP with 29% less computation with *11%* parameters *(Table 1)*.

Table 1. Comparison between YOLOv4 and YOLOR [7]

Model	**Size**	**FPS**$^{TitanRTX}_{batch32}$	**FLOPs**	**# parameters**	**AP**^{val}	**AP**$^{val}_{50}$	**AP**$^{val}_{75}$	**AP**$^{val}_{S}$	**AP**$^{val}_{M}$	**AP**$^{val}_{L}$
YOLOR-P6	1280	72	326.2G	37265016	52.5%	70.6%	57.4%	37.4%	57.3%	65.2%
YOLOR-W6	1280	66	454.0G	79873400	54.0%	72.1%	59.1%	38.1%	58.8%	67.0%
YOLOR-E6	1280	39	684.0G	115909400	54.6%	72.5%	59.8%	39.9%	59.0%	67.9%
YOLOR-D6	1280	31	936.8G	151782680	55.4%	73.5%	60.6%	40.4%	60.1%	68.7%
Y4-P6	1280	34	718.4G	127530352	54.4%	72.7%	59.5%	39.5%	58.9%	67.3%
U5R5-S6	1280	139	69.6G	12653596	43.3%	61.9%	47.7%	29.0%	48.0%	53.3%
U5R5-M6	1280	93	209.6G	35889612	50.5%	68.7%	55.2%	35.5%	55.2%	62.0%
U5R5-L6	1280	67	470.8G	77218620	53.4%	71.1%	58.3%	38.2%	58.4%	65.7%
U5R5-X6	1280	36	891.6G	141755500	54.4%	72.0%	59.1%	40.1%	59.0%	67.2%

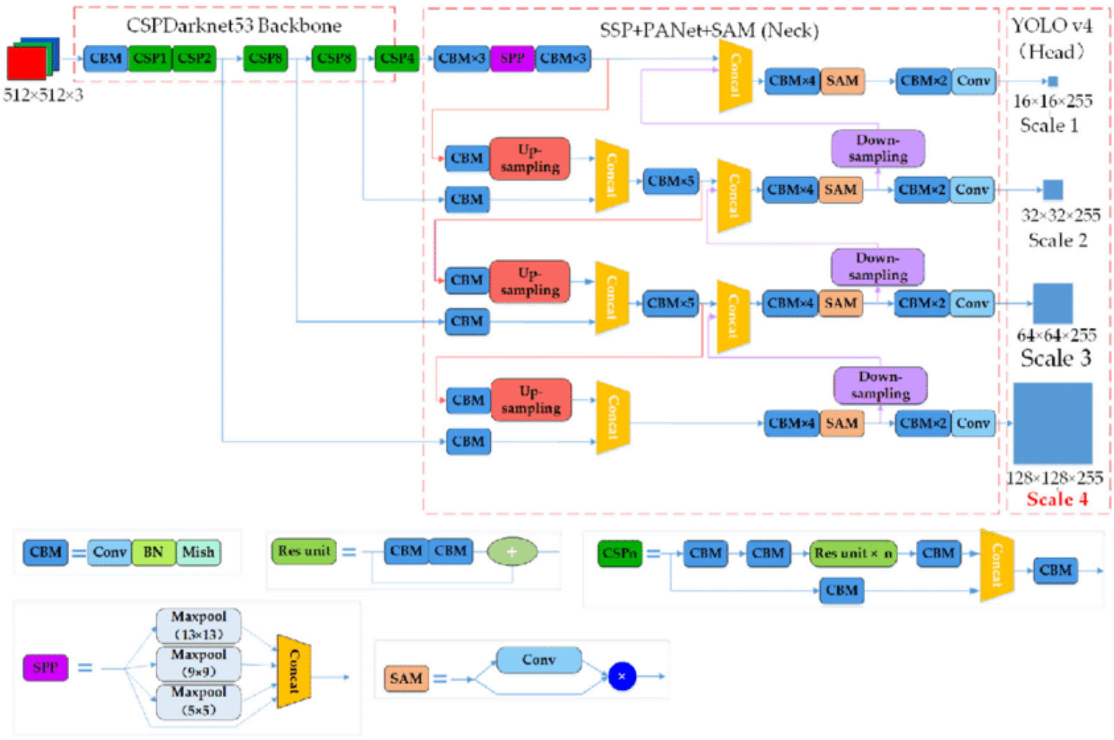


Figure 4. YOLO-R Architecture *[6]*

2.3. YOLO-X

After YOLOv5, advanced label assignment and anchor-free detectors grabbed the spotlight due to their better results. However, the YOLO family wasn't still integrated with these improvements which became the ultimate reason for the development of YOLOX model.

With a DarkNet53 backbone, YOLOX is a single-stage object detector that makes a number of changes to YOLOv3. In particular, YOLO's head is replaced with a decoupled one. Prior to adding two parallel branches with two 3 x 3 conv layers each for the classification and regression tasks, we adopt a 1 x 1 conv layer for each level of FPN features to minimize the feature channel to 256 [8].

To solve the conflict between classification and regression in object detection, this model uses Decoupled head architecture which means that classification and localization processes on the images are separated. This decoupled head architecture can be implemented on both regression and classification based models, one-stage and two-stage respectively.

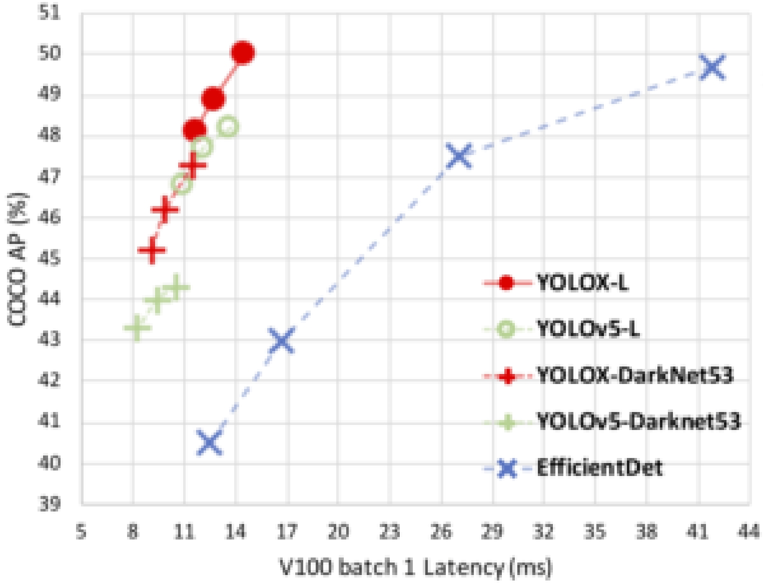


Figure 5. Comparison between YOLOX and YOLOv5 *[8]*

3. Method Used

In this paper, the YOLOv7 trained on MS COCO model was utilized on a soccer game play, detecting all the players in real-time and football movement. The inference time and accuracy of the model was carefully observed and then was compared with other models.

3.1. Major Improvements in Yolov7

YOLOv7 is the latest state of the art detector to this date among all known object detectors. Few architectural reforms have been introduced to the model that have improved the speed and the accuracy significantly. Just similar to scaled YOLOv4, this model also uses COCO dataset. The architectural reforms include E-ELAN (Extended Efficient Layer Aggregation Network) along with Model Scaling for concatenation based Models to enhance learning ability. Additionally, planned re-parameterized convolution and coarse for auxiliary with Fine for lead loss are aliased as the bag of freebies that also compliment the model for better learning without actually increasing the training cost [1].

Expand, shuffle, and merge cardinality is used in the proposed E-ELAN to continuously improve the network's capacity for learning while preserving the original gradient path. This block in YOLOv7 is based on the previously developed ELAN computational block *(Figure 6).*

Compound model scaling is another reason for the improved performance of YOLOv7. To make it more computing devices friendly, the model scaling was used to find the accuracy and speed requirements.

Usually for model scaling, *NAS* (Network Architecture Search) is used because of the parameter-specific scaling. However, in YOLOv7, the model is further optimized with a compound model scaling approach. In this approach, with and depth are coherently scaled for concatenation-based models *(Figure 7)*.

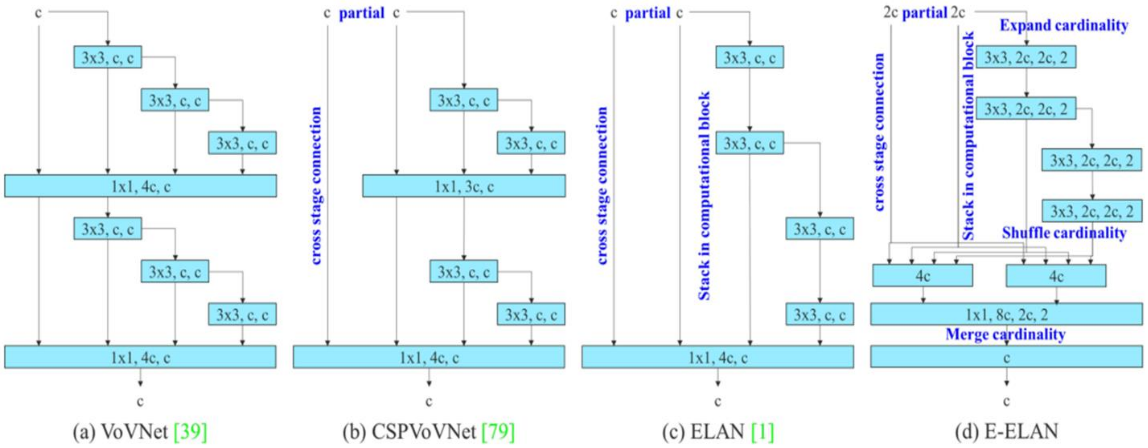


Figure 6. Comparison between VoVNET *[9]*, CSPVoVNET *[10]*, LAN and ELAN for better Network Learning

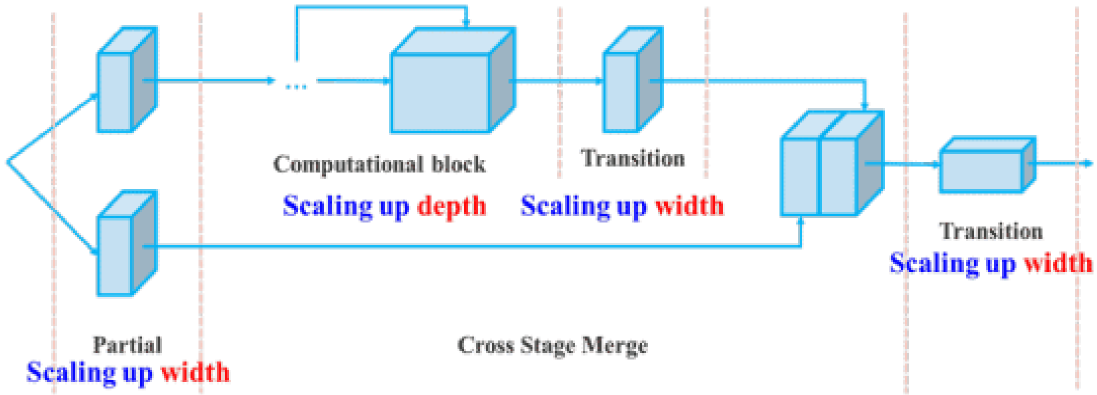


Figure 7. Compound Model Scaling in YOLOv7 *[1]*

Additionally, this model put forward the planned re-parameterization convolution that affects the accuracy significantly. Re-parameterization techniques [[11], [12], [13], [14], [15], [16]] are the method for enhancing the model after training. It lengthens the training process but yields better inference outcomes. Both Model level and Module level ensemble re-parametrization are the two types of re-parametrization used to finalize this models.

Model level re-parametrization is usually done either by training the multiple models with same settings but different training data and then taking the average of their weights or taking the average of the weights of models at multiple epochs. Recently, research on module level re-parameterization has exploded. This approach divides the model training process into a number of modules. To create the final model, the outputs are ensemble. *(Figure 8)*

The YOLO architecture usually contains a neck, a head and backbone and the outputs are obtained in the head. YOLOv7 has made few changes here as well. It is not constrained by a single head. It contains multiple heads to achieve whatever it wants. Earlier this method has also been used in Deep Supervision technique that is used by DL models to utilize multi heads.

There are two heads that. Firstly, the lead head in YOLOv7 is referred to as the final output head and secondly, the Auxiliary Head is the head that helps with middle-layer training. With the assistance of am assistant loss, auxiliary heads' weights are updated that allows the Deep Supervision [17] and the model's learning ability gets uplifted. This approach is closely related to Lead Head and the Label Assigner. Label Assigner is a method that assigns soft labels after taking the ground truth and network prediction outcomes into account. It's vital to notice that the label assigner creates coarse and soft labels rather than hard labels *(Figure 9)*.

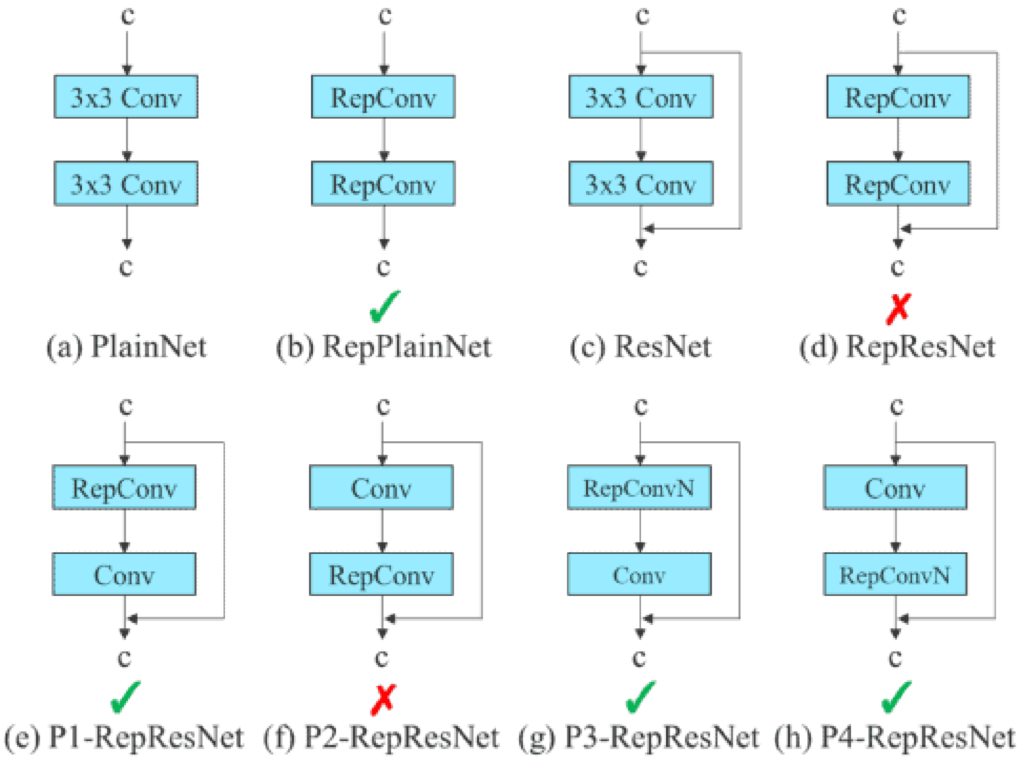


Figure 8. Best ways to perform module-level ensemble *[18]*

The YOLOv7 network's Lead Head makes predictions about the outcome. These final results are used to generate soft labels. The crucial aspect is that the identical soft labels that are generated are used to calculate the loss for both the lead head and the auxiliary head. In the end, the soft labels are used to train both heads. This is necessary because the lead head has a somewhat robust learning potential, which means that the soft label that results from it should be more accurate in capturing the distribution and correlation between the source and target data. The lead head will be better able to concentrate on learning residual information that has not yet been learnt by enabling the shallower auxiliary head immediately study the information that the lead head has learned.

Getting to the labels that go from coarse to fine *(Figure 9)*. Actually, two distinct soft labels sets are produced in the technique outlined above. The first one is fine label set for the lead head to train and a set of coarse labels for the auxiliary head's training. The fine labels are identical to the soft labels that were created immediately. More grids are, however, handled as positive targets in order to construct the coarse labels. To do this, the positive sample assignment procedure's limitations are relaxed.

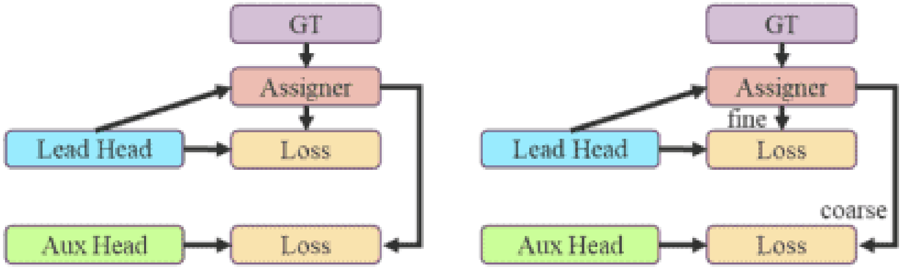


Figure 9. Lead guided Assigner and Coarse to fine Lead Guided Assigner

YOLOv7 demonstrates *56.8%* AP against *28 ms* inference time as compared to YOLO-R that exhibits *56.4%* AP against same inference time. All other models including YOLOv5, Scaled YOLOv4, PPYOLOE and YOLOX. All these models demonstrate lower *AP* as compared to the new YOLOv7 *(Figure 10)*.

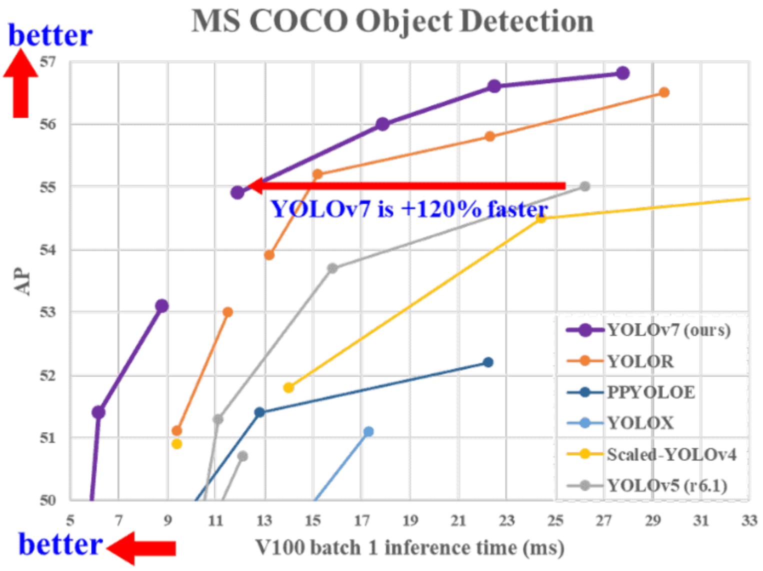


Figure 10. Comparison between YOLOv7 and other models of YOLO Family [1]

4. Results And Comparison Between YoloV5, Yolo-X, Yolo-R And YoloV7

After applying YOLOv7 along with YOLOR, YOLOv5 and YOLOX on the data, the performance of all these models in terms of *FPS* and *accuracy* was examined.

4.1. Fps Comparison

When *YOLOv5m* was applied to the data it showed the speed on *27* Frames per second which is followed by *YOLOR-p6* that yielded in *23 FPS*. *YOLOXm* also gave the *FPS* of 22 that is slightly less than that of *YOLOR-p6*. Finally, lowest speed was displayed by *YOLOv7* that gives the *FPS* of *17* only. *(Figure 11)* Indeed, YOLOv5m is delivering the highest FPS rate.

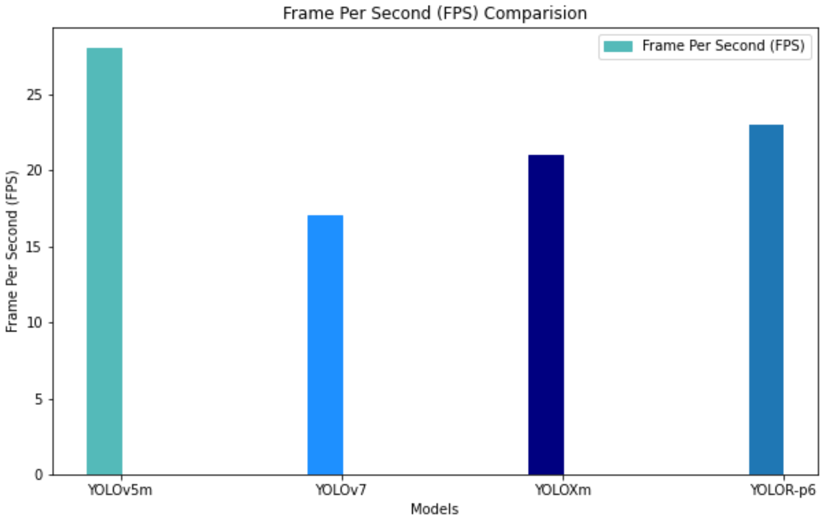


Figure 11. FPS Comparison between YOLOv5m, YOLO-p6, YOLOXm and YOLOv7

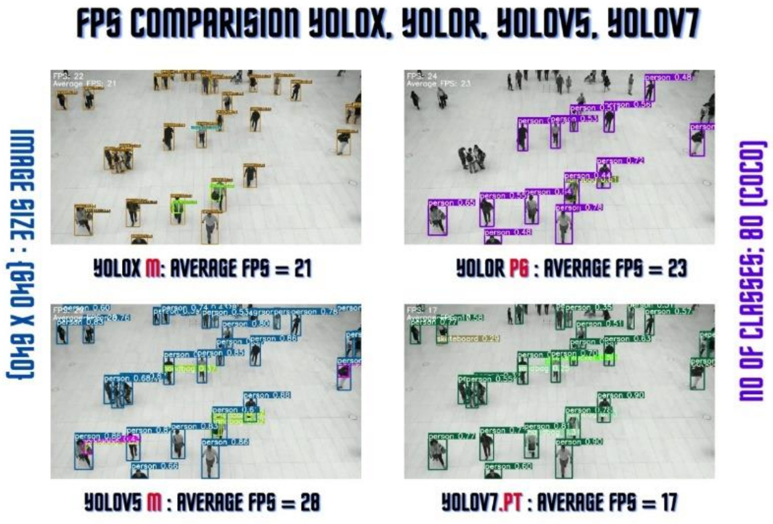

FPS COMPARISION YOLOX, YOLOR, YOLOV5, YOLOV7
IMAGE SIZE : {640 X 640}
YOLOX M: AVERAGE FPS = 21
YOLOR P6 : AVERAGE FPS = 23
YOLOV5 M : AVERAGE FPS = 28
YOLOV7.PT : AVERAGE FPS = 17
NO OF CLASSES: 80 {COCO}

Figure 12. FPS Comparison

4.2. Accuracy Comparison

The accuracy results of all these models were also compared and it was seen that even with the least speed, *YOLOv7* is delivering close to state of the art accuracy. Highest accuracy of *YOLORv7* can be seen in the metrics which is closely followed by the *YOLO-p6* with the accuracy of *56% AP*. Then comes the *YOLOv5m* with the accuracy of *53% AP* along with *YOLOXm* that gave the *46% AP (Figure 13)*.

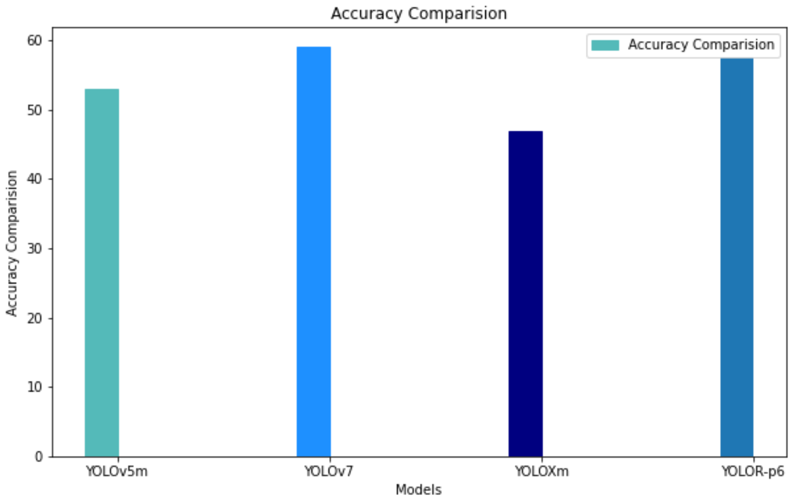


Figure 13. Accuracy Comparison between YOLOv5m, YOLOR-p6, YOLOXm and YOLOv7

5. FUTURE WORK

Further improvements are expected to result in an improved accuracy and faster frames per second rate while detecting players of both teams distinctly. This method could be refined more by also considering to track players and their movements which can be a great help in detecting off-sides too on real time, as offside detection is still one of the most challenging decisions in soccer at the moment, a goal ruled out wrongly because of error in offside detection by referee can change the game result. Furthermore, an application can be built on top of this model to distinguish the substitute players from the spectators so their movements can also be tracked closely. Semantics and instance segmentation can be used for that. The performance analysis of the models in this paper are being compared after training them on MS COCO dataset. These models can be trained on the custom data and then the FPS comparison between them can be observed. These custom trained model can be implemented on the embedded devices like Jetson Nano for various uses. The change in the FPS in those devices and the use of quantization techniques to deal with the FPS drop is yet another discussion. Additionally, a user friendly dashboard to interpret the information about game trend, insight and likelihood of goal by teams can be created.

6. CONCLUSION

Over the past few years, YOLO family has seen some tremendous research work as it has been proved to be a great resource for real-time object detection. Due to its fast & accurate detection, YOLO algorithm has huge potential for being used in the commercial applications. In this paper, different variants of the YOLO algorithms are analyzed for the accuracy and speed to yield the best performing model. YOLOv5, YOLO-X, YOLO-R and YOLOv7 models were applied to a soccer game play video to track the moving ball and the players in the playground. All the models performed well however YOLOv7 proved to be state of the art in terms of the accuracy showing *58% AP* whereas YOLOR-p6 which comes right after the former model and shows *56% AP*. On the other hand, YOLOv7 performs very poorly in terms of FPS demonstrating only 17 frames per second coming last against YOLO-R, YOLO-X and YOLOv5. Our work proves that where YOLOv7 is not the fastest on our data however due to its accuracy, it can be used in commercial applications like User friendly Soccer analyzer dashboard for insight analysis and likelihood of goal. In edge devices, there is a limitation of using YOLOv7 as we need at least 5 FPS for certain applications however the small variant YOLOv7.pt can only provide 3 to 4 FPS that are not enough. After some modifications, it can also be used in embedded devices to support applications like retail store monitoring and autonomous robots & vehicles.

REFERENCES

[1] A. B. H.-Y. M. L. Chien-Yao Wang, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors".

[2] S. D. R. G. A. F. Joseph Redmon, "You Only Look Once: Unified, Real-Time Object Detection".

[3] G. C. J. W. X. Y. J. H. Xingxing Xie, "Oriented R-CNN for Object Detection".

[4] A. F. Joseph Redmon, "YOLO9000: Better, Faster, Stronger".

[5] H. L. K. L. L. C. Y. L. Renjie Xu, "A Forest Fire Detection System Based on Ensemble Learning".

[6] S. Z. C. Z. R. L. Xiqi Wang, "R-YOLO: A Real-Time Text Detector for Natural Scenes with Arbitrary Rotation".

[7] I.-H. Y. a. H.-Y. M. L. Chien-Yao Wang, "You Only Learn One Representation: Unified Network for Multiple Tasks".

[8] S. L. F. W. Z. L. J. S. Zheng Ge, "YOLOX: Exceeding YOLO Series in 2021".

[9] J.-w. H. S. L. Y. B. J. P. Youngwan Lee, An Energy and GPU-Computation Efficient Backbone Network for Real-Time Object Detection.

[10] A. B. H.-Y. M. L. Chien-Yao Wang, Scaled-YOLOv4: Scaling Cross Stage Partial Network.

[11] V. V. S. I. J. S. a. Z. W. Christian Szegedy, Rethinking the Inception Architecture for Computer Vision.

[12] Y. L. G. P. Z. L. J. E. H. K. Q. W. Gao Huang, Snapshot Ensembles: Train 1, get M for free.

[13] A. T. a. H. Valpola, Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results.

[14] P. I. D. P. D. P. V. a. A. G. W. Timur Garipov, Loss surfaces, mode connectivity, and fast ensembling of DNNs.

[15] D. P. T. G. D. V. a. A. G. W. Pavel Izmailov, Averaging weights leads to wider optima and better generalization.

[16] X. Z. Y. Z. J. H. G. D. a. J. S. Xiaohan Ding, Scaling up your kernels to 31x31: Revisiting large kernel design in CNNs.

[17] S. X. P. G. Z. Z. a. Z. T. Chen-Yu Lee, Deeply-supervised nets.

[18] X. Z. N. M. J. H. G. D. J. S. Xiaohan Ding, RepVGG: Making VGG-style ConvNets Great Again.

AUTHORS

Ismat Saira Gillani is a graduate student of AI and Robotic Engineering at Columbus State University, USA. She is passionate about Computer Vision, Robotics and Machine Learning. She is always ready to adapt new situation, solve problems efficiently and achieve productivity goals. She focuses on making use of SoTA algorithms to build applications and design customized AI pipelines to provide the best possible AI solutions and generate measurable business value.

Muhammad Rizwan Munawar has received his BS in Computer Science from the COMSATS University Islamabad, Pakistan, currently he is a computer vision Engineer in a Multi-National company. He is obsessed with Machine Vision and Embedded Systems, focusing on the development of vision base applications, refining model architectures for research tasks, and developing custom modules that are suitable for his research.

Muhammad Talha is a Electrical and Telecom Engineer from Pakistan. He is passionate about Computer Vision, Machine Learning, Deep Learning, Point Cloud Processing and Video Compression. Currently, he is pursuing PhD from University of Missouri Kansas City, United States.

Muhammad Salman Azhar, Double Gold Medalist is Computer Science graduate from COMSATS Pakistan, is currently doing his Master's in AI from Air University Islamabad, Pakistan. He is an expert advisor of Financial Markets. Love building Trading Bots. He is a skilled Developer and passionate about AI, VR and Robotics.

Yousra Mashkoor is a Software Engineer, She's an active OpenSource contributor and is one of the top 110 Github contributors from Pakistan. Yousra is an ardent individual who's passionate about emerging high tech and how it is revolutionizing. She has expertise in Blockchain, Cloud Architecture, and Data Science.

Muhammad Sami uddin is highly passionate in Computer Vision and Deep Learning. He has built numerous Computer vision-based projects. Enthusiastic learner with mathematical background. Currently working on 3D computer vision models.

Usama Zafar is a Software engineer from University of Faisalabad, Pakistan. He loves to code and solve problems. He is passionate about Computer Vision, Machine learning and Artificial Intelligence. He is obsessed with Computer vision and Deep Learning.

Wireless Secret Sharing Game between Two Legitimate Users and an Eavesdropper

Lei Miao, Hongbo Zhang, and Dingde Jiang

Dept. of Engineering Technology, Middle Tennessee State University,
Murfreesboro, TN 37132, USA
Dept. of Engineering Technology, Middle Tennessee State University,
Murfreesboro, TN 37132, USA
School of Astronautics & Aeronautic, University of Electronic Science and
Technology of China, Sichuan, China

Abstract. Wireless secret sharing is crucial to information security in the era of Internet of Things. One method is to utilize the effect of the randomness of the wireless channel in the data link layer to generate the common secret between two legitimate users Alice and Bob. This paper studies this secret sharing mechanism from the perspective of game theory. In particular, we formulate a non-cooperative zero-sum game between the legitimate users and an eavesdropper Eve. In a symmetrical game where Eve has the same probability of successfully receiving a packet from Alice and Bob when the transmission distance is the same, we show that both pure and mixed strategy Nash equilibria exist. In an asymmetric game where Eve has different probabilities of successfully receiving a packet from Alice and Bob, a pure strategy may not exist; in this case, we show how a mixed strategy Nash equilibrium can be found.

Keywords: secret sharing, wireless communications, game theory, Nash equilibrium.

1 Introduction

Security and privacy in wireless networking relies on symmetric-key cryptography, which requires pre-established private keys at both the transmitter and the receiver. In the era of Internet of Things (IoT) where Machine to Machine (M2M) communications frequently occur with minimum human intervention, automatic and secure sharing of secrets for the purpose of cryptography is crucial to information security. There are various ways to share secrets automatically in wireless networks. One direction is to combine cryptographic schemes and channel coding techniques so that transmitted messages between two legitimate users Alice and Bob cannot be decoded by the eavesdropper Eve [1] [2]. Recent works along this line can be found in [3], [4], and [5] for interference, broadcast, and multiple access channels, respectively. Another approach exploits the principle of reciprocity [6] in wireless communications and extracts the secret from the common observation between Alice and Bob on the wireless channel state [7] [8] [9]. All these methods mentioned above are collectively known as the physical layer solutions, which essentially exploit the

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 29-39, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121603

randomness and varying nature of wireless channels to share secrets. They do not work very well when the speed of variation in wireless channels is slow and may also require costly modifications to existing communication protocols and infrastructure.

In a different direction, the effect of wireless channel dynamics on the data link layer is utilized to share secrets [10], [11], [12]. The idea behind works along this line is as follows: Alice and Bob keep sending each other unicast packets without retry, using which the secret is derived; Eve would eventually lose a packet and be unable to figure out the secret even if she knew exactly the mechanism Alice and Bob use. More details of this approach can be found in our previous work [13] where we discuss optimal secret sharing between Alice and Bob with the presence of Eve. Specifically, we assume in [13] that Eve's location is random, and only Alice and Bob can choose how to generate the secret; we show that when the probability of successfully transmitting a packet is monotonically decreasing with the transmission distance and Eve's location is uniformly distributed, the optimal strategy for Alice and Bob to minimize the probability that Eve figures out the secret is to generate half of the secret from each one of them.

In this paper, we consider the case that Eve can also choose her location in order to maximize her probability of receiving all packets and figuring out the secret. Specifically, we assume that both the legitimate users (Alice and Bob) and the eavesdropper (Eve) do not know each other's strategy but are both rational. Let P_e be the probability of Eve figuring out the secret. Then, Alice and Bob's goal is to minimize P_e or maximize $-P_e$, and Eve's goal is to maximize P_e. This observation motivates us to formulate the problem as a zero-sum game between the legitimate users and the eavesdropper.

Security games have been studied extensively on the interaction between legitimate and malicious users, and game-theoretic approaches have been applied to a wide range of problems, including security at the physical and MAC layers, security at the application layer, cryptography, etc. For comprehensive reviews, see [14] [15] [16]. Our secret sharing game is different from the existing ones in the literature: we study how to share secrets using the effect of the unreliable nature of wireless channels on the data link layer. Our results are based on the probability function of Eve successfully receiving a packet. Nonetheless, our analysis does not rely on a specific form of the probability function; instead, our work would be applicable to any probability function as long as a mild assumption is satisfied. The main contributions of this paper is as follows: *(i)* We show that the optimal secret sharing problem can be considered as a game between two legitimate users and the eavesdropper; *(ii)* We analyze the symmetric game case and identify both pure and mixed strategy Nash equilibria; *(iii)* For the asymmetric game case, we discover two different scenarios that yield pure and mixed Nash equilibrium, respectively; and *(iv)* We show how the mixed strategy Nash equilibrium can be found when the probabilities of successful packet transmission are known.

The organization of the rest of the paper is as follows: in Section 2, we discuss the system model and formulate the game; in Section 3, we present the main results of the optimal secret sharing zero-sum game; and finally, we conclude in Section 4.

2 System Model and Problem Formulation

In our system model, the two legitimate users Alice and Bob are at two different locations that are D meters away, and they are trying to exchange N packets $\{Pkt_1, Pkt_2, \ldots, Pkt_N\}$, using which the secret is calculated. One simply way to obtain the secret is to exclusive-OR all N packets together: $secret = Pkt_1 \oplus Pkt_2 \oplus \cdots \oplus Pkt_N$. Due to the unreliable nature of wireless communications, Eve will have high probability of losing one or more packets when N is large so that she will not be able to figure out the secret. Without loss of generality, we let N be an even number. For ease of notation, We assume that each of the two game players, i.e., the legitimate users and the eavesdropper, has three strategies. For Alice and Bob, the three strategies are: Alice sends all N packets to Bob, Bob sends all N packets to Alice, and each one of them sends $N/2$ packets to the other. We use S_A, S_B, and S_{AB} to denote these three strategies, respectively. Eve chooses to stay somewhere between Alice and Bob, and she also has three different strategies: stay close to Alice, stay close to Bob, and stay in the exact middle. We use L_A, L_B, and L_M to denote these three locations/strategies, respectively. Note that although we only have three strategies defined for each player, our results can be extended to the cases that more strategies are available. We further assume that locations L_A and L_B are ϵ, $\epsilon \in (0, \frac{D}{2})$, meters away from Alice and Bob, respectively; location L_M is $\frac{D}{2}$ meters away from both Alice and Bob. Thus, $P_A(\epsilon)$, $P_A(D-\epsilon)$, and $P_A(\frac{D}{2})$ are the probabilities of Eve successfully receiving a packet from Alice when Eve's strategy is L_A, L_B, and L_M, respectively. Similarly, $P_B(\epsilon)$, $P_B(D-\epsilon)$, $P_B(\frac{D}{2})$ are the probabilities of Eve successfully receiving a packet from Bob when Eve's strategy is L_B, L_A, and L_M, respectively.

Let $P_A(d)$ and $P_B(d)$ be the probability of Eve successfully receiving a packet from Alice or Bob, respectively, when the transmission distance is d. We have the following assumption about $P_A(d)$ and $P_B(d)$.

(i) Each packet transmission is independent from each other; *(ii)* $P_A(d)$ and $P_B(d)$ are time-invariant; *(iii)* $P_A(\epsilon) > P_A(\frac{D}{2}) > P_A(D-\epsilon)$ and $P_B(\epsilon) > P_B(\frac{D}{2}) > P_B(D-\epsilon)$; and *(iv)* $P_A(\frac{D}{2}) > \frac{1}{2}[P_A(D-\epsilon) + P_A(\epsilon)]$ and $P_B(\frac{D}{2}) > \frac{1}{2}[P_B(D-\epsilon) + P_B(\epsilon)]$. The assumptions above is quite generic and does not require the exact form of functions $P_A(d)$ and $P_B(d)$. Parts *(i)* and *(ii)* above are valid in slow-fading environments where the coherence time of the wireless channel is long and the channel state is stable during the period of secret sharing. Part *(iii)* states that the key factor that determines the probability of successful packet transmission is the distance, which is especially true in long-distance wireless communications. An example of $P_A(d)$ and $P_B(d)$ supporting the monotonicity assumption in VANET

(Vehicular Ad Hoc Networks) environments can be found in [17], in which Killat et al. simulate and verify a theoretical probability of successful transmission function of distance inferred from the Nakagami-m distribution of RF wave propagation. It is well known that in free space, the path loss of RF signals is proportional to the square of distance; part *(iv)* above reflects this: in spite of random factors such as channel fading, the signal's power and the probability of successful transmission attenuates faster when the distance is larger.

3 Optimal Secret Sharing as a Zero-Sum Game

Let s_L and s_E be the strategies of the legitimate users, i.e., Alice and Bob, and Eve, the eavesdropper, respectively. We have $s_L \in \{S_A, S_B, S_{AB}\}$ and $s_E \in \{L_A, L_B, L_M\}$ We use $U_L(s_L, s_E) = -P_e$ and $U_E(s_L, s_E) = P_e$ to denote the utility functions of the legitimate users and Eve, respectively. Essentially, Alice and Bob would like to minimize the probability of Eve figuring out the secret, and Eve would like to maximize the same probability.

Definition 1. *A strategy profile (s_L^*, s_E^*) is a Nash equilibrium if $U_L(s_L^*, s_E^*) \geq U_L(s_L, s_E^*)$ for each feasible strategy s_L and $U_E(s_L^*, s_E^*) \geq U_E(s_L^*, s_E)$ for each feasible strategy s_E.*

3.1 Symmetric Game

We first consider a symmetric game scenario that the following hold:

$$P_A(L_A) = P_B(L_B) = P(\epsilon),\ P_A(L_B) = P_B(L_A) = P(D - \epsilon),$$
$$\text{and } P_A(L_M) = P_B(L_M) = P(D/2).$$

We have the utility matrix shown in Table 1 where the utility functions of Eve are positive and the ones of Alice and Bob are negative. Next, let us first introduce an auxiliary lemma.

Table 1. Utility matrix of the symmeric game.

			Alice and Bob	
		S_A, q_1	S_{AB}, q_2	$S_B, 1 - q_1 - q_2$
	L_A, p_1	$\pm P^N(\epsilon)$	$\pm P^{\frac{N}{2}}(\epsilon) P^{\frac{N}{2}}(D - \epsilon)$	$\pm P^N(D - \epsilon)$
Eve	L_M, p_2	$\pm P^N(\frac{D}{2})$	$\pm P^N(\frac{D}{2})$	$\pm P^N(\frac{D}{2})$
	$L_B, 1 - p_1 - p_2$	$\pm P^N(D - \epsilon)$	$\pm P^{\frac{N}{2}}(\epsilon) P^{\frac{N}{2}}(D - \epsilon)$	$\pm P^N(\epsilon)$

Lemma 1. $P^{\frac{N}{2}}(\epsilon)P^{\frac{N}{2}}(D-\epsilon) < P^N(\frac{D}{2})$

Proof: Because $P(\epsilon) \in (0,1)$, $P(D-\epsilon) \in (0,1)$, and $P(\frac{D}{2}) \in (0,1)$, we only need to show that $P(\epsilon)P(D-\epsilon) < P^2(\frac{D}{2})$. Because $\epsilon \in (0, \frac{D}{2})$, $D-\epsilon \neq \epsilon$. Since $P(\cdot)$ is monotonically decreasing, we have

$$[P(D-\epsilon) - P(\epsilon)]^2 = P^2(D-\epsilon) + P^2(\epsilon) - 2P(D-\epsilon)P(\epsilon) > 0,$$

i.e.,

$$\frac{1}{4}[P^2(D-\epsilon) + P^2(\epsilon)] > \frac{1}{2}[P(D-\epsilon)P(\epsilon)]. \tag{1}$$

From part *(iv)* of Assumption 2, we have

$$\begin{aligned} P^2(\frac{D}{2}) &= P^2(\frac{1}{2}(D-\epsilon) + \frac{1}{2}\epsilon) > [\frac{1}{2}P(D-\epsilon) + \frac{1}{2}P(\epsilon)]^2 \\ &= \frac{1}{4}[P^2(D-\epsilon) + P^2(\epsilon)] + \frac{1}{2}[P(D-\epsilon)P(\epsilon)] \end{aligned}$$

Invoking (1), we have $P^2(\frac{D}{2}) > P(D-\epsilon)P(\epsilon)$ ■.

We are now ready to discuss the pure strategy result of the symmetric game.

Lemma 2. *Strategy profile (S_{AB}, L_M) is a pure strategy Nash equilibrium.*

Proof: It can be seen from the utility matrix that $U_L(S_{AB}, L_M) = U_L(S_A, L_M) = U_L(S_B, L_M) = -P^N(\frac{D}{2})$. Invoking Lemma 1, we have

$$U_E(S_{AB}, L_M) = P^N(\frac{D}{2}) > P^{\frac{N}{2}}(\epsilon)P^{\frac{N}{2}}(D-\epsilon) = U_E(S_{AB}, L_A) = U_E(S_{AB}, L_B).$$

From Definition 1, it follows that strategy profile (S_{AB}, L_M) is a pure strategy Nash equilibrium. ■

Lemma 2 indicates that in the pure strategy Nash equilibrium, Alice and Bob each generates half of the packets and Eve stays in the middle location L_M. We now turn our attention to a mixed strategy Nash equilibrium, in which Eve has probabilities p_1, p_2, and $p_3 = 1 - p_1 - p_2$ to use strategies L_A, L_M, and L_B, respectively; similarly, Alice and Bob have probabilities q_1, q_2, and $q_3 = 1 - q_1 - q_2$ to use strategies S_A, S_{AB}, and S_B, respectively.

Lemma 3. *In a mixed strategy Nash equilibrium, Eve's strategy is to stay at L_M with probability 1; Alice and Bob should have positive probabilities on all three strategies S_A, S_B, and S_{AB} so that:*

$$q_1P^N(\epsilon) + q_2P^{\frac{N}{2}}(\epsilon)P^{\frac{N}{2}}(D-\epsilon) + q_3P^N(D-\epsilon) < P^N(\frac{D}{2}) \tag{2}$$

and

$$q_1P^N(D-\epsilon) + q_2P^{\frac{N}{2}}(D-\epsilon)P^{\frac{N}{2}}(\epsilon) + q_3P^N(\epsilon) < P^N(\frac{D}{2}) \tag{3}$$

Proof: Suppose that $0 < q_1 < 1$, $0 < q_2 < 1$, and $0 < 1 - q_1 - q_2 < 1$. In a mixed strategy Nash equilibrium, we have:

$$\begin{aligned}
& - p_1 P^N(\epsilon) - p_2 P^N(\frac{D}{2}) - (1 - p_1 - p_2) P^N(D - \epsilon) \\
& = -p_1 P^{\frac{N}{2}}(\epsilon) P^{\frac{N}{2}}(D - \epsilon) - p_2 P^N(\frac{D}{2}) - (1 - p_1 - p_2) P^{\frac{N}{2}}(\epsilon) P^{\frac{N}{2}}(D - \epsilon) \\
& = -p_1 P^N(D - \epsilon) - p_2 P^N(\frac{D}{2}) - (1 - p_1 - p_2) P^N(\epsilon)
\end{aligned}$$

Solving the above equations, we get $p_1 = p_3 = 0$, and $p_2 = 1$. If it is the case in the mixed strategy Nash equilibrium, we must also have (2) and (3).

Next, we verify that when (2) and (3) hold, $\exists\ q_1, q_2$, and q_3 so that $0 < q_1 < 1$, $0 < q_2 < 1$, and $0 < q_3 < 1$. Let $q_1 = q_3$, and (2) and (3) become one inequality:

$$2q_1[P^N(\epsilon) + P^N(D - \epsilon)] + q_2 P^{\frac{N}{2}}(\epsilon) P^{\frac{N}{2}}(D - \epsilon) < P^N(\frac{D}{2}) = 2q_1 P^N(\frac{D}{2}) + q_2 P^N(\frac{D}{2}) \tag{4}$$

Invoking Lemma 1, we have $q_2 P^{\frac{N}{2}}(\epsilon) P^{\frac{N}{2}}(D - \epsilon) < q_2 P^N(\frac{D}{2})$. We now consider two cases.

Case 1: $2q_1[P^N(\epsilon) + P^N(D - \epsilon)] \leq 2q_1 P^N(\frac{D}{2})$. In this case, (4) always holds as long as q_1, q_2, and q_3 are nonzero probabilities.

Case 2: $2q_1[P^N(\epsilon) + P^N(D - \epsilon)] > 2q_1 P^N(\frac{D}{2})$. In this case, we can always pick small enough positive q_1 and q_3 values so that (4) holds. ■

3.2 Asymmetric Game

We now consider an asymmetric game scenario that $P_A(d) > P_B(d)$, i.e., when the transmission distance is the same, Eve has higher probability to successfully receive a packet from Alice than from Bob. For example, if Alice has higher transmission power than Bob or Bob is closer to a noise source, then the signal to noise ratio between Alice and Eve may be higher than that between Bob and Eve, causing the asymmetric game scenario described above. We have the following utility matrix shown in Table 2.

Table 2. Utility matrix of the asymmeric game.

		Alice and Bob		
		S_A, q_1	S_{AB}, q_2	$S_B, 1 - q_1 - q_2$
Eve	L_A, p_1	$\pm P_A^N(\epsilon)$	$\pm P_A^{\frac{N}{2}}(\epsilon) P_B^{\frac{N}{2}}(D - \epsilon)$	$\pm P_B^N(D - \epsilon)$
	L_M, p_2	$\pm P_A^N(\frac{D}{2})$	$\pm P_A^{\frac{N}{2}}(\frac{D}{2}) P_B^{\frac{N}{2}}(\frac{D}{2})$	$\pm P_B^N(\frac{D}{2})$
	$L_B, 1 - p_1 - p_2$	$\pm P_A^N(D - \epsilon)$	$\pm P_B^{\frac{N}{2}}(\epsilon) P_A^{\frac{N}{2}}(D - \epsilon)$	$\pm P_B^N(\epsilon)$

Note that similar to the utility matrix in the symmetric game case, we only show the utility functions of Eve; the ones of Alice and Bob are negative and are not shown above.

Lemma 4. *If $P_A(d) > P_B(d)$, and $P_B(\epsilon) \leq P_A(D-\epsilon)$, then strategy profile (S_B, L_B) is a pure strategy Nash equilibrium.*

Proof: Because $\epsilon \in (0, D/2)$ and $P_B(d)$ is monotonically decreasing, we have

$$P_B^N(\epsilon) > P_B^N(\frac{D}{2}) > P_B^N(D-\epsilon), \text{i.e.,}$$

$$U_E(S_B, L_B) > U_E(S_B, L_M) > U_E(S_B, L_A). \tag{5}$$

By assumption, $P_B(\epsilon) \leq P_A(D-\epsilon)$, we get

$$P_B^N(\epsilon) \leq P_B^{\frac{N}{2}}(\epsilon) P_A^{\frac{N}{2}}(D-\epsilon) \leq P_A^N(D-\epsilon). \tag{6}$$

Multiplying (6) by -1 yields:

$$-P_B^N(\epsilon) \geq -P_B^{\frac{N}{2}}(\epsilon) P_A^{\frac{N}{2}}(D-\epsilon) \geq -P_A^N(D-\epsilon), \text{i.e.,}$$

$$U_L(S_B, L_B) \geq U_L(S_{AB}, L_B) \geq U_L(S_A, L_B). \tag{7}$$

Combining (5) and (7), it follows that strategy profile (S_B, L_B) is a pure strategy Nash equilibrium. ■

The intuition behind Lemma 4 is that if $P_B(d)$ is so much less than $P_A(d)$ so that $P_B(\epsilon) \leq P_A(D-\epsilon)$, then the best strategy of the legitimate users is to always let Bob send the packets; conversely, the best strategy of Eve is to stay close to Bob so that she could maximize the probability of receiving all packets.

Lemma 5. *If $P_A(d) > P_B(d)$ and $P_B(\epsilon) > P_A(D-\epsilon)$, then there is no pure strategy Nash equilibrium.*

Proof: We discuss the three columns of the utility matrix individually.
(1) Column #1: We have $U_E(S_A, L_A) = P_A^N(\epsilon) > U_E(S_A, L_M) = P_A^N(\frac{D}{2}) > U_E(S_A, L_B) = P_A^N(D-\epsilon)$ due to part *(iii)* of Assumption 2. Therefore, only strategy profile (S_A, L_A) can possibly be a pure strategy in the first column. However, we have $U_L(S_A, L_A) = -P_A^N(\epsilon) < -P_B^N(\epsilon) < -P_B^N(D-\epsilon) = U_L(S_B, L_A)$ in the first row. Therefore, there is no pure strategy Nash equilibrium in the first column of the utility matrix.

(2) Column #2: Because $U_E(S_A, L_A) = P_A^N(\epsilon) > U_E(S_{AB}, L_A) = P_A^{\frac{N}{2}}(\epsilon) P_B^{\frac{N}{2}}(D-\epsilon) > U_E(S_B, L_A) = P_B^N(D-\epsilon)$, strategy profile (S_{AB}, L_A) cannot be a pure strategy Nash equilibrium. Similarly, because $U_E(S_A, L_M) = P_A^N(\frac{D}{2}) > U_E(S_{AB}, L_M) = P_A^{\frac{N}{2}}(\frac{D}{2}) P_B^{\frac{N}{2}}(\frac{D}{2}) > U_E(S_B, L_M) = P_B^N(\frac{D}{2})$, strategy profile (S_{AB}, L_M) cannot be

a pure strategy Nash equilibrium either. Finally, because $U_E(S_A, L_B) = P_A^N(D - \epsilon) < U_E(S_{AB}, L_B) = P_A^{\frac{N}{2}}(D-\epsilon)P_B^{\frac{N}{2}}(\epsilon) < U_E(S_B, L_B) = P_B^N(\epsilon)$, strategy profile (S_{AB}, L_B) cannot be a pure strategy Nash equilibrium.
(3) Column #3: Similarly to the Column #1 case, there is no pure strategy Nash equilibrium in the third column either. The analysis is very similar to the Column #1 case, and we omit the details. ■

Lemma 5 shows that when $P_B(\epsilon) > P_A(D - \epsilon)$, i.e., $P_B(d)$ is not too much less than $P_A(d)$, no pure strategy Nash equilibrium exists. According to [18], at least one mixed strategy Nash equilibrium always exists in this case. The utility functions are:

$$-p_1P_A^N(\epsilon) - p_2P_A^N(\frac{D}{2}) - p_3P_A^N(D-\epsilon) \qquad (q_1)$$

$$-p_1P_A^{\frac{N}{2}}(\epsilon)P_B^{\frac{N}{2}}(D-\epsilon) - p_2P_A^{\frac{N}{2}}(\frac{D}{2})P_B^{\frac{N}{2}}(\frac{D}{2}) - p_3P_B^{\frac{N}{2}}(\epsilon)P_A^{\frac{N}{2}}(D-\epsilon) \qquad (q_2)$$

$$-p_1P_B^N(D-\epsilon) - p_2P_B^N(\frac{D}{2}) - p_3P_B^N(\epsilon \qquad (q_3)$$

$$q_1P_A^N(\epsilon) + q_2P_A^{\frac{N}{2}}(\epsilon)P_B^{\frac{N}{2}}(D-\epsilon) + q_3P_B^N(D-\epsilon) \qquad (p_1)$$

$$q_1P_A^N(\frac{D}{2}) + q_2P_A^{\frac{N}{2}}(\frac{D}{2})P_B^{\frac{N}{2}}(\frac{D}{2}) + q_3P_B^N(\frac{D}{2}) \qquad (p_2)$$

$$q_1P_A^N(D-\epsilon) + q_2P_B^{\frac{N}{2}}(\epsilon)P_A^{\frac{N}{2}}(D-\epsilon) + q_3P_B^N(\epsilon) \qquad (p_3)$$

where (q_1), (q_2), and (q_3) are the payoffs of the legitimate users when strategies S_A, S_{AB}, and S_B are used, respectively; (p_1), (p_2), and (p_3) are the payoffs of Eve when strategies L_A, L_M, and L_B are used, respectively. The procedure of finding the mixed strategy Nash equilibrium involves two steps: *proposition* and *verification*. In the first step, we make an assumption about either $\{p_1, p_2, p_3\}$ or $\{q_1, q_2, q_3\}$ and use the utilization functions to solve for the other set of probabilities. If the solution is feasible and we are able to use it in the second step to verify that the proposition provided in Step 1 is indeed true, the Nash equilibrium is found. Next, we formally present the procedure in Algorithm 1 where we only show the propositions about $\{p_1, p_2, p_3\}$; the pseudo code of making propositions about $\{q_1, q_2, q_3\}$ is very similar.

3.3 Numerical Example

In this subsection, we present a numerical example. For ease of calculation, we let $N = 2$. The probabilities are: $P_A(\epsilon) = 0.99$, $P_A(\frac{D}{2}) = 0.94$, $P_A(D - \epsilon) = 0.80$, $P_B(\epsilon) = 0.90$, $P_B(\frac{D}{2}) = 0.84$, and $P_B(D-\epsilon) = 0.70$. Invoking Lemma 5, there is no pure strategy Nash equilibrium. The mixed strategy utility functions corresponding to (q_1) through (p_3) are:

$$-0.3401p_1 - 0.2364p_2 - 0.64 \qquad (8)$$

Algorithm 1 Finding mixed strategy Nash equilibrium in an asymmetric game when $P_B(\epsilon) > P_A(D - \epsilon)$

```
Proposition: enumerate the following assumptions.
$p_1, p_2$, and $p_3$ are all non-zero probabilities; solve $(p_1)$=$(p_2)$=$(p_3)$ for $q_1, q_2$, $q_3$ and go to Verification.
For any two probabilities $p'$ and $p'' \in \{\mathrm{p}_1, p_2, p_3\}$, assume they are non-zero and use $p'''$ to denote the remaining probability. Solve $(p')$=$(p'')$>$(p''')$ for $q_1, q_2, q_3$ and go to Verification.
Verification:
if the solution is infeasible then
    Continue to the next assumption
else
    if $q_1$,$q_2$, and $q_3$ are all positive then
        Solve $(q_1)$=$(q_2)$=$(q_3)$ for $p_1, p_2, p_3$.
    end if
    if two probabilities $q'$ and $q'' \in \{\mathrm{q}_1, q_2, q_3\}$ are positive, and the remaining probability $q'''$ is 0 then
        Solve $(q') = (q'') > (q''')$ for $p_1, p_2, p_3$.
    end if
    if $q' \in \{\mathrm{q}_1, q_2, q_3\}$ is 1, and the other two probability $q''$ and $q'''$ are 0 then
        Solve $(q') > (q'')$ and $(q') > (q''')$ for $p_1, p_2, p_3$.
    end if
    if the solution of $p_1, p_2, p_3$ matches with the proposition then
        Nash equilibrium is found and exit
    else
        Continue to the next assumption
    end if
end if
```

$$0.027p_1 - 0.0504p_2 - 0.72 \tag{9}$$

$$0.32p_1 + 0.1044p_2 - 0.81 \tag{10}$$

$$0.4901q_1 + 0.203q_2 + 0.49 \tag{11}$$

$$0.178q_1 + 0.084q_2 + 0.7056 \tag{12}$$

$$-0.17q_1 - 0.09q_2 + 0.81 \tag{13}$$

We start out by assuming that $p_1 \in (0,1)$, $p_2 \in (0,1)$, and $1 - p_1 - p_2 \in (0,1)$. Under this proposition, we have (11) = (12) = (13), whose solution is $q_1 = 1.946$, $q_2 = -3.292$, and $1 - q_1 - q_2 = 2.346$. This is infeasible, meaning that p_1, p_2, and $1 - p_1 - p_2$ cannot be all positive and less than 1.

Next, we discuss three cases of p_1, p_2, and $1 - p_1 - p_2$.

Case 1: $p_1 \in (0,1)$, $p_2 \in (0,1)$, and $1 - p_1 - p_2 = 0$. It yields that (11) = (12) > (13). There are two solutions that ensure q_1, q_2, and $1 - q_1 - q_2$ are not all positive and less than 1. Therefore, we have two subcases:

Case 1.1: $q_1 = 0.6908$, $q_2 = 0$, and $1 - q_1 - q_2 = 0.3092$. It implies that at equilibrium, we must have (8) = (10) > (9), which has no solution between 0 and 1 for p_1 and p_2.

Case 1.2: $q_1 = 0.5469$, $q_2 = 0.4531$, and $1 - q_1 - q_2 = 0$. It implies that at equilibrium, we must have (8) = (9) > (10), which has solutions of p_1 and p_2 so that $p_1 + p_2 \in (0,1)$. It implies that $1 - p_1 - p_2 \in (0,1)$, which is impossible.

Case 2: $p_1 \in (0,1)$, $p_2 = 0$, and $1 - p_1 - p_2 \in (0,1)$. It yields that (11) = (13) > (12). There are no solutions.

Case 3: $p_1 = 0$, $p_2 \in (0,1)$, and $1 - p_1 - p_2 \in (0,1)$. In this last case, we have (12) = (13) > (11). The only feasible solution to it is $q_1 = 0$, $q_2 = 0.6$, and $1 - q_1 - q_2 = 0.4$. If this solution is also the one in equilibrium, we need to have (9) = (10) > (8), which also has a feasible solution: $p_1 = 0$, $p_2 = 0.5814$, and $1 - p_1 - p_2 = 0.4186$.

It completes the numerical example, and the mixed Nash equilibrium is as follows:

$$\begin{aligned}(p_1, p_2, 1 - p_1 - p_2) &= (0, 0.5814, 0.4186)\\ (q_1, q_2, 1 - q_1 - q_2) &= (0, 0.6, 0.4)\end{aligned}$$

4 Conclusions

We have studied the optimal secret sharing problem between two legitimate users (Alice and Bob) and an eavesdropper (Eve), formulated as a non-cooperative zero-sum game. In the symmetric game case, both pure and mixed strategy Nash equilibria exist. Our results indicate that regardless of the type of the equilibrium, Eve should always stay in the middle of Alice and Bob. In the pure strategy Nash equilibrium, the best strategy of Alice and Bob is to generate half of the packets from each one of them; in a mixed strategy Nash equilibrium, Alice and Bob could generate all the packets from one user only, but some inequalities involving the probabilities must hold.

In the asymmetric game case that Eve has better chance to successfully receive packets from Alice than from Bob, we show that there are two scenarios: if it is very asymmetrical, then a pure strategy Nash equilibrium exists, in which Bob is the one who generates all the packets and Eve chooses to stay near Bob; o.w., a mixed strategy equilibrium exists and can be calculated.

References

1. Barros J, Rodrigues MRD. Secrecy Capacity of Wireless Channels. In: ; 2006; Seattle, WA.
2. Maurer UM. Secret Key Agreement by Public Discussion from Common Information. *IEEE Trans. on Information Theory* 1993; 39: 733-742.
3. Chen J. Secure communication over interference channel: To jam or not to jam?. *IEEE Transactions on Information Theory* 2019; 66(5): 2819–2841.
4. Hyadi A, Rezki Z, Alouini MS. Securing Multi-User Broadcast Wiretap Channels with Finite CSI Feedback. *IEEE Transactions on Information Theory* 2020.
5. Mukherjee P, Ulukus S. Secure degrees of freedom of the multiple access wiretap channel with multiple antennas. *IEEE Transactions on Information Theory* 2018; 64(3): 2093–2103.
6. Balanis CA. *Antenna Theory: Analysis and Design.* New York: John Wiley and Sons. 2nd ed. 1997.
7. Mathur S, Trappe W, Mandayam N, Ye C, Reznik A. Radio-telepathy: extracting a secret key from an unauthenticated wireless channel. In: ; 2008; San Francisco, CA, USA.
8. Miao L. Differential Secret Sharing in Wireless Networks. *IEEE Wireless Communications Letters* 2015; 4(2): 213-216.
9. Ruotsalainen H, Zhang J, Grebeniuk S. Experimental Investigation on Wireless Key Generation for Low-Power Wide-Area Networks. *IEEE Internet of Things Journal* 2019; 7(3): 1745–1755.
10. Xiao S, Gong W, Towsley D. Secure Wireless Communication with Dynamic Secrets. In: ; 2010; San Diego, CA.
11. Yao T, Fukui K, Nakashima J, Nakai T. Initial common secret key sharing using random plaintexts for short-range wireless communications. *IEEE Trans. on Consumer Electronics* 2009; 55: 2025-2033.
12. Safaka I, Fragouli C, Argyraki K, Diggavi S. Creating shared secrets out of thin air. In: ACM. ; 2012: 73–78.
13. Miao L, Jiang D. Optimal secret sharing for wireless information security in the era of Internet of Things. *Personal and Ubiquitous Computing* 2019: 1–16.
14. Manshaei MH, Zhu Q, Alpcan T, Bacşar T, Hubaux JP. Game theory meets network security and privacy. *ACM Computing Surveys (CSUR)* 2013; 45(3): 1–39.
15. Abdalzaher MS, Seddik K, Elsabrouty M, Muta O, Furukawa H, Abdel-Rahman A. Game theory meets wireless sensor networks security requirements and threats mitigation: A survey. *Sensors* 2016; 16(7): 1003.
16. Do CT, Tran NH, Hong C, et al. Game theory for cyber security and privacy. *ACM Computing Surveys (CSUR)* 2017; 50(2): 1–37.
17. Killat M, Hartenstein H. An empirical model for probability of packet reception in vehicular ad hoc networks. *EURASIP Journal on Wireless Communications and Networking* 2009; 2009(721301): 12.
18. Nash J. Non-cooperative games. *Annals of mathematics* 1951: 286–295.

MIMICKING A COMPLETE LIFE-CYCLE OF FIAT CURRENCY IN ONE E-CASH SYSTEM

Peifang Ni[1, 2]

[1]TCA Laboratory, Institute of Software,
Chinese Academy of Sciences Beijing, China
[2]State Key Laboratory of Cryptology, Beijing, China

ABSTRACT

The electronic cash was introduced by Chaum in 1982 and now many e-cash systems have been proposed in order to mimic the fiat currency. Bitcoin provides us with an attractive way to construct a decentralized e-cash system. Ideally, we would like to make the system more practical, for example, the users can be able to transfer coins between each other multiple times and they can also withdraw arbitrary amount of coins rather than one or the predefined number, so that in the spend protocol the user can spend any amount of valid coins.

In this paper, we propose a provably secure and more practical e-cash system. Firstly, it can provide the anonymous transfer of coins between users, so that the merchant can spend the received coins further; secondly, the user can withdraw arbitrary amount of coins rather than the one or predefined number; thirdly, during the transfer of coins, the coins have a fixed size; finally, the fair exchange between the users can also be achieved.

KEYWORDS

E-cash, fiat currency, coin, anonymous, preventing double-Spending, fair exchange.

1. INTRODUCTION

This document describes, and is written to conform to, author guidelines for the journals of AIRCC series. It is prepared in Microsoft Word as a .doc document. Although other means of preparation are acceptable, final, camera-ready versions must conform to this layout. Microsoft Word terminology is used where appropriate in this document. Although formatting instructions may often appear daunting, the simplest approach is to use this template and insert headings and text into it as appropriate.

Bitcoin [1] is the most prominent cryptocurrencies, whose security does not rely on any single trusted third party and the transaction ledger is publicly available. But now the cash is still the most prevalent payment method in real life because that cash transaction is anonymous, transferable and it can prevent the double-spending attack. Furthermore, there exists no unfair exchange when the users exchange with physical cash. Here we focus on the construction of a more practical e-cash system [2] that can simulate the physical currency better. It is known that the proof of coins' validity is based on some personal message, which is easy to break the users' privacy. Moreover, in the e-cash system, the user withdraws coins from the bank then sends it to the merchant who must deposits it to the bank rather than spends it further. And in the trustless network, the unfairness occurs between users, e.g., the malicious payer or payee can refuse to sending the valid messages.

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 41-52, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121604

Obviously, the ideal e-cash system should be equipped with the same security properties of the physical cash. There has been number of works aim to solve these problems, but these works usually solve several problems and also bring some new problems, for example, using *malleable signature* to achieve the transitivity of coins [3], so the users can get a valid signature of another message according to some transformation without communicating with the bank, but this scheme brings another problem that the size of the coin is increasing during the process of transfer.

1.1. Our Contributions

In this paper, we present a provably secure and more practical e-cash system that is closet to the physical cash。In this system, the only trusted component is the blockchain and we achieve that:

- **transitivity:** the valid coins can be used further by the payee;
- **arbitrary amount of coins:** the user can withdraw arbitrary amount of valid coins from the bank;
- **prevent double-spending:**any user cannot spend a same coin twice;
- **fair exchange:** none of users can cheat in the trustless network.

1.2. Related Work

A complete e-cash system should satisfies: anonymity, transitivity, divisibility, the ability to prevent double-spending attacks and fair exchange among users. Now we describe the related works in the following aspects.

Transitivity means that, in e-cash system, the merchant (or payee) can spend the received coin further without depositing it to the bank firstly. Okamoto and Oha[4,5] are the first to propose transferable e-cash, but they only provided the weak anonymity. Then in 1992 [6] proved that the size of transferable coins is increasing during the transfer process. Baldimtsi [3] showed us the first transferable e-cash system that satisfies all of the anonymity properties and provides us with a new efficient double-spending detection mechanism, but the size of coins is increasing during the transfer process. The works [7,8] proposed a fully anonymous and transferable e-cash scheme that satisfies all the security properties, which can prevent double-spending attack with the help of blockchain. However, the size of coins is still increasing during the transfer process, and it presents a noneffective solution for this problem that, when the size reaches the upper bound, the user can deposit it to the bank and then the bank declares that this coin is invalid.

Divisibility means that the user can withdraw a unique coin of value $\leq 2^n$ from the bank and spend it in several times to some distinct merchants. Compact e-cash allows the user to withdraw a wallet from the bank that contains 2^n coins in the withdraw protocol [9,10], however, this scheme only allow the user to spend one coin each time, which makes the scheme unpractical. Canard proposed the first efficient divisible e-cash system secure in the standard model [11]. And Tewari showed us the e-cash system that the user can withdraw arbitrary denomination of coin [7], which must be spent in one time.

Fair exchange between users means that, in the trustless network, the malicious user may refuse returning the corresponding message when he learns some knowledge. Constructing contracts on blockchain provides us with a feasible method [12,13] and the works [14,15,16] show us how bitcoin can be used in the area of *MPC,* where the fairness means that the dishonest users will pay a fine for the honest ones as a compensation. In the e-cash system the participants may face the following problem that shall the buyer pays the merchant first then the merchant sends the buyer service or the other way around ? Is there a way that none of the parties can cheat the other one? The works [17,18] used blockchain to achieve the fair exchanges among users.

1.3. Organization

In section 2, we review some classical definitions and notations. Section 3 describes the Bitcoin simply. In section 4, we present our constructions of the e-cash system. And the security proof and conclusion are showed in section 5 and 6 respectively.

2. PRELIMINARIES

2.1. Assumptions

Our construction is based on the following assumptions:

- The only ``trusted component" is the Blockchain that each user has access to it (denoted by $Ledger$), and if a transaction is on the $Ledger$ then it means that the transaction is provably valid. And the communication channel between users and $Ledger$ is secure;
- The amount of coins that the user can withdraw is limited;
- The coin has $l \geq 1$different denominations (to correspond with the physical cash, it can have other definitions)$Val_1 > Val_2 > \cdots Val_l$ and the bank holds the corresponding signature key pairs denoted as $(pk_1, sk_1), \dots, (pk_l, sk_l)$;
- Each denomination has a fixed prefix and the form is $\mathrm{Value, serial\ number, signature}$, i.e., $\mathrm{Coin}_1 = (\mathrm{Val}_1,\ \mathrm{SN}_1, \sigma_1)$ means that the denomination of Coin_1 is Val_1, its serial number is SN_1 and it can be provably valid with σ_1 and pk_1;

2.2. E-Cash system

As the traditional e-cash system [19,20], our scheme consists of two basic parties, the user U (or the merchant M) and bank B, and the following algorithms:

- $ParamGen(1^\lambda)$: on input security parameter λ, the parameter generation algorithm outputs the system public parameters par. We assume that par is the default input to the remaining algorithms;
- $KeyGen((1^\lambda))$: the key generation algorithm is executed by U and B respectively, and outputs (pk_U, sk_U) and (pk_B, sk_B) ,where $(pk_B, sk_B) =: \left((pk'_B, sk'_B), (pk''_B, sk''_B)\right) = ((pk'_B, sk'_B); (pk_1, sk_1), \dots, (pk_l, sk_l))$;
- $Register(B[pk_B, sk_B], U[pk_U, sk_U])$: the registration protocol allows U to have an account in the bank and, at the end, both parties output ok for success or $\perp$ for the failure of registration;
- $Withdraw(B[pk''_B, sk''_B], U[pk_U, sk_U])$: the withdraw protocol allows U to get the coins with value Val, which can be provable with the fixed prefix and the signature of B;
- $Spend(U_1[coin, pk_{U_1}, sk_{U_1}], U_2[coin, pk_{U_2}, sk_{U_2}])$: the spend protocol allows user U_1 with $coin$ to buy the U_2's service ω;
- $Deposit(U[coin, pk_U, sk_U], B[pk_B, sk_B])$: the deposit protocol allows the user U to deposit $coin$ to his account held by$B$.

2.3. Digital Signature Scheme

A digital signature scheme consists of a triple of probabilistic polynomial-time algorithms $(Gen, Sign, Ver)$ satisfying the followings:

- Key-generator algorithm Gen: on input 1^λ, Gen outputs (pk, sk);
- Signing algorithm $Sign$: on input secret key sk and a message $\alpha \in \{0,1\}^*$, $Sign$ outputs signature σ;
- (Deterministic) Verifying algorithm Ver: on input public key pk, message $\alpha \in \{0,1\}^*$ and signature σ, Ver returns $output \ \in \{accept, reject\}$;
- For each key pair (pk, sk) in the range of$Gen(1^\lambda)$ and message $\alpha \in \{0,1\}^*$, the algorithms $Sign$ and Ver satisfy

$$\Pr\left[Ver\big(pk, \alpha, Sign(sk, \alpha)\big)\right] = 1$$

where the probability is taken over the internal coin tosses of $Sign$ and Ver.

2.4. Public Encryption Scheme

A public encryption scheme consists of a triple of probabilistic polynomial-time algorithms $(EncGen, Enc, Dec)$ satisfying the followings:

- Key-generator algorithm $EncGen$: on input 1^λ, $EncGen$ outputs (sk, pk);
- Encryption algorithm Enc: on input public key pk and message $\alpha \in \{0,1\}^*$, Enc outputs a ciphertext $c \coloneqq Enc_{pk}(\alpha)$;
- Decryption algorithm Dec: On input private key sk, a ciphertext c, Dec outputs $\alpha' = Dec_{sk}(c)$.
- For each key pair (sk, pk) in the range of $EncGen(1^\lambda)$ and message $\alpha \in \{0,1\}^*$, the algorithms Enc and Dec satisfy:

$$\Pr\left[Dec_{sk}\left(Enc_{pk}(\alpha)\right) = \alpha\right] = 1$$

where the probability is taken over the internal coin tosses of algorithms Enc and Dec.

2.5. Zero-knowledge Proofs of Knowledge

In the $spend$ protocol, the user will use the ZKPoK to let the merchant believe that the $coin$ exactly belongs to him and the merchant uses $ZKPoK$ to let the user believe that he knows the service/witness ω.

A pair of interactive Turing machines $< P, V >$ is called an interactive proof system for a language L if machine V is polynomial-time and the following two conditions hold:

- Completeness: there exists a negligible function c such that for every $x \in L$,

$$\Pr[< P, V > (x) = 1] > 1 - c(|x|)$$

- Soundness: there exists a negligible function s such that for every $x \notin L$,

$$\Pr[< P, V > (x) = 1] < s(|x|)$$

$c(\cdot)$is called the completeness error and $s(\cdot)$ is the soundness error.

Zero-knowledge protocol is the interactive proof system with zero-knowledge property, which means that the prover can convince the verifier that some instance $x \in L$ without providing the verifier with any additional information beyond the fact.

A zero-knowledge protocol is called a zero-knowledge proof of knowledge if $L \in NP$ and for every prover P^* there exists a polynomial-time machine, called knowledge extractor, that can interact withP^*, and at the end it outputs x. We follow the requirement in [16], without loss of generality, the last two messages in the protocol are challenge x sent by the verifier and prover's response r. The extractor extracts x after being given transcripts of two accepting executions.

2.6. Security Properties

In this section we give the security properties by redefining these of [3]. Firstly, we show the $Oracles$ that the adversary $\mathcal{A}$ can interact in the security definitions.

- $Create(1^\lambda)$ executes $(pk_U, sk_U) \leftarrow$ UKeyGen() and outputs pk_U;
- $BRegister(pk_U)$ plays the bank side in the $Register$ protocol and interacts with $\mathcal{A}$. If pk_U has been in the bank database, then abort; otherwise, $\mathcal{A}$ owns an account in bank B;
- $BWithdraw$ plays the bank side in the $Withdraw$ protocol and interacts with $\mathcal{A}$;
- $S\&R$ is the $Spend\ and\ Receive$ oracle that allows $\mathcal{A}$ to observe the process of $Spend$ protocol between honest users;
- $UReceive$ plays the $merchant$ side in the $Spend$ protocol and interacts with $\mathcal{A}$;
- $USpend$ plays the $user$ side in the $Spend$ protocol and interacts with $\mathcal{A}$;
- $UDeposit$ plays the $user$ side in the $Deposit$ protocol and interacts with $\mathcal{A}$.

Unforgeability means that the adversary $\mathcal{A}$ should not be able to forge a valid coin (the signature of bank) without communicating with bank or forges a valid coin according to the coins that he has owned so that to spend more coins than he has withdrew from the bank. We use the following experiment to define ***Unforgeability***.

Experiment $\boldsymbol{Expt}_{\mathcal{A}}^{unforg}(\lambda)$

- $par \leftarrow ParamGen\ (1^\lambda)$;
- $\sigma \leftarrow \mathcal{A}^{\text{Creat,BRegister,BWith}}(par)$;
- let qW, qD be the amount of coins that $\mathcal{A}$ calls to $BWithdraw, BDeposit$ respectively.

If $(Ver_{pk_B}(\sigma) = 1) \wedge (qW < qD)$, then return 1; otherwise, return 0.

Definition 1 (***Unforgeability***) An E-Cash system is unforgeable if for any probabilistic polynomial-time adversary $\mathcal{A}$, we have the $\boldsymbol{Expt}_{\mathcal{A}}^{unforg}(\lambda)$ defines as：

$$\Pr\left[\boldsymbol{Expt}_{\mathcal{A}}^{unforg}(\lambda) = 1\right] < negl(\lambda)$$

Preventing double − spending means that no user is able to spend a transaction twice. We use the following experiment to define ***Preventing double − spending***.

Experiment $\boldsymbol{Expt}_{\mathcal{A}}^{\text{PreDS}}(\lambda)$

- $par \leftarrow ParamGen\ (1^{\lambda})$;
- $\mathcal{A}$ broadcasts two transactions to redeem the same transaction;

If these two transactions are confirmed in the $Ledger$, then return 1; otherwise, return 0.

Definition 2 (***Preventing double − spending***)An E-Cash system can prevent double-spending if for any probabilistic polynomial-time adversary $\mathcal{A}$, we have the $\boldsymbol{Expt}_{\mathcal{A}}^{\text{PreDS}}(\lambda)$ defines as：

$$\Pr\left[\boldsymbol{Expt}_{\mathcal{A}}^{\text{PreDS}}(\lambda) = 1\right] < negl(\lambda)$$

Anonymity. We define anonymity in three aspects as $\{Observe - then - Receive$, $Spend - then - Observe$ and $Spend - then - Receive$ [3] respectively.

$Observe - then - Receive$ means that the adversary $\mathcal{A}$ cannot link a coin that he receives as a bank during the $Deposit$ protocol or user during the $Spend$ protocol to a coin that he observed transitivity between two honest users before.

Experiment $\boldsymbol{Expt}_{\mathcal{A},b}^{\text{OtR}}(\lambda)$

- $par \leftarrow ParamGen\ (1^{\lambda})$;
- $(coin_0, \text{coin}_1, \delta) \leftarrow \mathcal{A}^{\text{Create,S \& }R}(par)$;
- if $\delta = 1$ then simulate $USpend(coin_b)\}$ to $\mathcal{A}$; otherwise, simulate $UDeposit(coin_b)\}$ to $\mathcal{A}$ ($b \in \{0,1\}$);
- $b^* \leftarrow \mathcal{A}^{\text{Create,S \& }R,USpend,UDeposit}(\text{par})$;

If $b^* = b$,then return 1; otherwise, return 0.

$Spend - then - Receive$ means that the adversary $\mathcal{A}$ cannot link a coin that he spent as a user during the $Spend$ protocol to a coin that he receives during the $Spend$ protocol as a merchant or during the $Deposit$ protocol as a bank.

Experiment $\boldsymbol{Expt}_{\mathcal{A},b}^{\text{StR}}(\lambda)$

- $par \leftarrow ParamGen\ (1^{\lambda})$;
- $(coin_0, \text{coin}_1, \delta) \leftarrow \mathcal{A}^{\text{Create,UReceive}}(par)$;
- if $\delta = 1$ then simulate $USpend(coin_b)\}$ to $\mathcal{A}$; otherwise, simulate $UDeposit(coin_b)\}$ to $\mathcal{A}$ ($b \in \{0,1\}$);
- $b^* \leftarrow \mathcal{A}^{\text{Create,S \& }R}(\text{par})$;

If $b^* = b$,then return 1; otherwise, return 0.

$Spend - then - Observe$ means that the adversary $\mathcal{A}$ can not link a coin that he spent as a user during the $Spend$ protocol to a coin that he observes the transitivity between two honest users.

Experiment $\boldsymbol{Expt}^{\mathrm{StO}}_{\mathcal{A},b}(\lambda)$

- $par \leftarrow ParamGen\ (1^{\lambda})$;
- $(coin_0, \mathrm{coin}_1) \leftarrow \mathcal{A}^{\mathrm{Create,UReceive}}(par)$;
- simulate $S\&R(coin_b)$;
- $b^* \leftarrow \mathcal{A}^{\mathrm{Create,UReceive,S\,\&\,}R}(\mathrm{par})$

If $b^* = b$,then return 1; otherwise, return 0.

Definition 3(***Anonymity***) An E-cash system is anonymous if for any probabilistic polynomial-time adversary $\mathcal{A}$, we have $\boldsymbol{Expt}^{\mathrm{OtR}}_{\mathcal{A},b}(\lambda), \boldsymbol{Expt}^{\mathrm{StR}}_{\mathcal{A},b}(\lambda)$ and $\boldsymbol{Expt}^{\mathrm{StO}}_{\mathcal{A},b}(\lambda)$ define as:

$$\Pr\left[\boldsymbol{Expt}^{\mathrm{OtR,StR,StO}}_{\mathcal{A},b}(\lambda) = 1\right] < negl(\lambda)$$

3. BITCOIN

Firstly, we recall the simplified version of bitcoin's transaction. Let $A = (A.pk, A.sk)$ be a key pair and the form of transaction that user A transfers the coin with value v to user B is as following:

$$T_x = (y, B.pk, v, \sigma)$$

where y is an index of the previous transaction $T_y(y = H(T_y)$ and $B.pk$ is the recipient of T_x, we also say that T_y is redeemed by T_x. The transaction T_x is valid if:

- $A.pk$ is the recipient of T_y;
- the value of T_y is at least v;
- transaction T_y has not been redeemed earlier;
- the signature σ of A is correct.

And there is also another condition that the transaction may have several ``inputs" and we do not use this form in our construction so we will not describe it in details. Furthermore, we describe a more detailed version. In the real-world bitcoin system, the user has more flexibility in defining the conditions that how the transaction can be redeemed. The transaction T_y contains a description of a function (output-script) π_y whose output is Boolean and the transaction T_x that can redeem T_y if π_y evaluates to true with input T_x. The transaction is defined as:

$$T_x = (y, \pi_x, v, \sigma, t)$$

Where $[T_x] = (y, \pi_x, v)$ is the $body$ of T_x and σ is the witness that is used to make π_y evaluates to true with input T_x. the scripts π_x are written in the Bitcoin scripting language.The transaction T_x is valid if:

- time t is reached;
- $\pi_y([T_x], \sigma)$ is true;
- transaction T_y has not been redeemed before.

In our construction, we define the transaction that is sent by bank B as $T = (\wedge, \pi, v, \sigma)$ (the index of the transaction that T redeems is empty) and we also define the transaction that user U wants to send to B to deposit the coins as $T = (H(T'), pk_B, \sigma)$, where T' is the transaction that be redeemed by T.

4. A MORE PRACTICAL E-CASH SYSTEM

In this section, we present the detailed construction of a more practical E-Cash system, which satisfies the properties of transitivity, anonymity, preventing double-spending, the size of coin is fixed during the transfer process.

For our construction we have the following assumptions: secure channels for all the communications so that an adversary cannot overhear or tamper with the transferred messages; in the $Spend$ protocol the user wants to buy the witness w from U_1, which satisfies $U's$ demand (i.e.: U_1 knows x that $f(x) = true$ which is harder to find x than verifying that $f(x) = true$ holds) with the value Val' and in the process that U proves knowledge of signatures and U_1 proves knowledge of the witness w will by executing a zero-knowledge proof of knowledge protocol with each other in the cut-and-choose technique, which is denoted as π.

Before giving the description of our construction, we show PoK of signature in details and PoK of witness w is similar.

- U_1 divides each signature $\sigma_{i,j}$ to n parts $\sigma_{i,j}^1, \dots, \sigma_{i,j}^n$ and each part is committed separately by computing $\tau_{i,j}^k := Commit(\sigma_{i,j}^k)$, where $i = 1, \dots, l, j \in \{1, \dots, m_1\} \cup \dots \cup \{1, \dots, m_l\}$and $k = 1, \dots, n$, and sends the commitments $\tau_{i,j}^k$ to U_2;
- U_2 sends the set $J \subseteq \{(i, j, k): i = 1, \dots, \mathrm{l}, j \in \{1, \dots, m_1\} \cup \dots \cup \{1, \dots, m_l\}, \mathrm{k} = 1, \dots, \mathrm{n}\}$, where the size of each J is smaller than n_0 so that U_2 cannot use any n_0-out-of-n parts to learn any knowledge of $\sigma_{i,j}$ andn_0 is predefined;
- U_1 opens the proper $\sigma_{i,j}$ according to subset J;
- U_2 verifies the openings.

Now, we present *a more practical E-Cash system*:

- $par \leftarrow ParamGen\left(1^\lambda\right)$: input the security parameter λ output the public parameter par and we assume that par is the default input to the remaining algorithms;
- $\{(pk_B, sk_B) \leftarrow BKeyGen(1^\lambda)$: the bank generates the key pairs for different denominations as $\left((pk'_B, sk'_B); (pk_1, sk_1), \dots, (pk_l, sk_l)\right) := (pk_B, sk_B)$:=($\left(pk'_B, sk'_B\right)$, (pk''_B, sk''_B));
- $\{(pk_U, sk_U) \leftarrow UKeyGen\left(1^\lambda\right)$: the user generates the key pairs as (pk_U, sk_U);
- Registration$(\mathrm{U}[(pk_U, sk_U), \mathrm{B}[(pk_B, sk_B)]])$: the user U sends pk_U to bank B, if $pk_U \in DB$ (the database of the bank) then B outputs $\perp$; otherwise, B selects a nonce $nonce$, computes $\mathrm{ID}_U = H(pk_U || nonce)$, adds (pk_U, ID_U) to DB and returns ID_U to U;
- $Withdraw(\mathrm{U}[(pk_U, sk_U), (pk'_U, sk'_U)], \mathrm{B}[(pk''_B, sk''_B)]])$: the user U selects a new key pair (pk'_U, sk'_U) and sends (pk'_U, ID_U,Val) where Val denotes the amount of coins that he will withdraw from B. B checks $U's$ account and its balance, returns $\perp$ if one of them is invalid; otherwise, B does the followings:
  - computes the number of coins with different denominations as $m_1, \dots, m_l$;
  - computes signatures for each coin: selects $s \in Z_p$ (where p is a prime) and computes $\sigma_{i,j} = Sign_{sk_i}(s || j || \mathrm{ID}_U)$ ($\sigma_{i,j}$ is the signature of the j^{th} coin wit denomination Val_i), where $i = 1, \dots, \mathrm{l}, j \in \{1, \dots, m_1\} \cup \dots \cup \{1, \dots, m_l\}$;

    - prepares the transactions $T_{i,j}^{U} = \left(\wedge, \pi_{i,j}, Val_i, \sigma_{i,j}^{B}\right)$, where $T_{i,j}^{U}$ means the j^{th} coin with denomination Val_i is transferred and $\pi_{i,j}\left(T'_{i,j}\right) = 1$ if $Verify_{pk'_U}\left(\left[T'_{i,j}, \sigma'_{i,j}, \left[T'_{i,j}\right]\right]\right) = 1$;
    - broadcasts transactions $T_{i,j}^{U}$ and sends $Coin = \left(\sigma_{1,1} \ldots \sigma_{1,m_1}, \ldots, \sigma_{l,1} \ldots \sigma_{l,m_l}\right)$ to user U;
    - deducts user $U's$ account.

- $\text{Spend}\left(U_1\left[\text{coin}, \left(pk'_{U_1}, sk'_{U_1}\right)\right], U_2\left[\omega, \left(pk_{U_2}, sk_{U_2}\right)\right)\right]$:first we assume that U_1 wants to buy a witness ω from U_2, whose value is Val', so in this process U_1 transfers Val' to U_2 and get witness ω:

    - U_1 computes the number of coins with different denominations as $m_1, \ldots, m_l$ and prepares the coins:

    $$coin_{i,j} = (Val_i, SN_i, \sigma_{i,j})$$

    where $i = 1, \ldots, \mathrm{l}, j \in \{1, \ldots, m_1\} \cup \ldots \cup \{1, \ldots, m_l\}$. And then U_1 computes $\pi = PoK(\sigma_{1,1} \ldots \sigma_{1,m_1}, \ldots, \sigma_{l,1} \ldots \sigma_{l,m_l})$ and sends $coin = (coin_{1,1}, \ldots, coin_{l,m_l}, \pi)$to U_2;
    - U_2verifies that U_1 owns $coin$ exactly, and then computes $h = H(\omega), \pi = PoK(\omega)$ and sends (h, π) to U_1;
    - U_1verifies (h, π) and prepares the following transactions:

    $$\text{PutMoney}_{i,j}^{U_1} = (H\left(T_{i,j}^{U_1}\right), \pi'_{i,j}, Val_i, \sigma_{sk'_{U_1}}, t)$$

    $$\text{ClaimMoney}_{i,j}^{U_1}(H\left(\text{PutMoney}_{i,j}^{U_1}\right), \pi_{i,j}^{U_1}, Val_i, \sigma_{sk'_{U_1}}, t')$$

    where$\pi'_{i,j} = ((\text{Verify}_{pk_{U_1}}\left(T_{i,j}\right) \wedge H(\omega) = h) \vee (\text{Verify}_{pk'_{U_1}}\left(T'_{i,j}\right) \wedge \pi_{i,j}^{U_1} = (\text{Verify}_{pk'_{U_1}}\left(T_{i,j}\right)))$;
    - U_2 prepares the following transactions:

    $$\text{ClaimMoney}_{i,j}^{U_2} = (H\left(\text{PutMoney}_{i,j}^{U_1}\right), \pi_{i,j}^{U_2}, Val_i, \sigma_{sk_{U_2}}, \omega, t_1)$$

    where $\pi_{i,j}^{U_2} = Verify_{pk_{U_2}}(T_{i,j})$;
    - U_1 broadcasts transactions $\text{PutMoney}_{1,1}^{U_1}, \ldots, \text{PutMoney}_{l,m_l}^{U_1}$ to the $Ledger$ and computes $c = \text{Enc}_{pk_{U_2}}\left(\sigma_{1,1}, \ldots, \sigma_{1,m_1}; \ldots; \sigma_{l,1}, \ldots, \sigma_{l,m_l}\right)$.If not all the transactions appear on the $Ledger$ or U_2 does not receive the correct c in time t, then the protocol halts;
    - U_2 waits for all the transactions $\text{PutMoney}_{1,1}^{U_1}, \ldots, \text{PutMoney}_{l,m_l}^{U_1}$ appearing on the $Ledger$,he broadcasts transactions $\text{ClaimMoney}_{1,1}^{U_2}, \ldots, \text{ClaimMoney}_{l,m_l}^{U_2}$ to the $Ledger$,which will reveal the witness ω. And if none of these transactions appears on the $Ledger$ in time t_1 , then U_1 broadcasts transactions $\text{ClaimMoney}_{1,1}^{U_1}, \ldots, \text{ClaimMoney}_{l,m_l}^{U_1}$to get his coins back.

- $Deposit(U\left[coin, \left(pk'_U, sk'_U\right)\right], B[pk_B, sk_B])$: before the $Deposit$ protocol starts, we assume that $Ledger$ contains transactions $T_i(i = 1, \ldots, n_l)$ (we also say that the number of coins that U owns is n_1) and each transaction can be redeemed by U. First U prepares the following transactions whose recipient is bank B (pk_B), then B adds the amount of coins to $U's$ account.

    - U prepares transactions $T'_1 = \left(H(T_1), pk_B, Val_1, \sigma_{[T'_1]}\right), \ldots, T'_{n_1} = \left(H\left(T_{n_l}\right), pk_B, Val_l, \sigma_{[T'_l]}\right)$ and broadcasts them to the $Ledger$, and sends $c = Enc_{pk_B}(\sigma_1, \ldots, \sigma_{n_l}, (pk_U, ID_U), (pk', sk'))$ to the bank B,where (pk', sk') is the key pair that signs the $T'_1, \dots, T'_n$, so that the adversary cannot disguise U to let the bank to add the coins to his account because he dose not know the correct key pair;

- Bwaits the transactions appear on the Ledgerand checks the received messages, if each of them is valid then adds the amount of coins to the $U's$ account, otherwise returns $\perp$.

5. Security Analysis

Theorem. Suppose the Digital Signature Scheme, Public Encryption Scheme, Zero-knowledge Proofs of Knowledge and the bitcoin blockchain protocol are secure, then our e-cash system satisfies unforgeability, anonymity, transferability, divisibility, preventing double-spending attacks and fair exchange between users.

Proof: ***(unforgeability)*** According to the definition of unforgeability (Section 2.6), we give the proof of unforgeability into two parts.

Part 1: Let the adversary $\mathcal{A}$ to execute $Create, Register$ and $Withdraw$ protocols with the $Oracle$, we denote it as simulator $\mathcal{S}$, without the knowledge of sk_B. By executing $Create$ and $Register$ protocols, the adversary $\mathcal{A}$ has an account in the bank's database DB. Then at the end of executing $Withdraw$ protocol with $\mathcal{S}$, the output of $\mathcal{S}$ is $coin = ((coin_1, \sigma_{1,1}, \dots, \sigma_{1,.m_1}), \dots, (coin_l, \sigma_{l,1}, \dots, \sigma_{l,.m_l}))$. With the security of Digital Signature Scheme that, without the knowledge of signature key sk_B, $\mathcal{S}$ can give a valid signature $\mathcal{S}$ with negligible probability.

Part 2: Let the adversary $\mathcal{A}$ to execute $Create, Register, Withdraw$ and $Deposit$ protocols with the $Oracle$, we denote it as simulator $\mathcal{S}$ with the knowledge of sk_B, and gets the valid coins $coin = ((coin_1, \sigma_{1,1}, \dots, \sigma_{1,.m_1}), \dots, (coin_l, \sigma_{l,1}, \dots, \sigma_{l,.m_l}))$ and the valid transactions that can be redeemed by $\mathcal{A}$. Thus if $\mathcal{A}$ succeeds at this game, then $\mathcal{A}$ must redeem the same transaction twice and both are confirmed on the $Ledger$ to accomplish the $Deposit$ protocol. With the security of Bitcoin system, $\mathcal{A}$ can redeem the same transaction more than once with negligible probability. So in this game, $\mathcal{A}$ succeeds with negligible probability.

(Preventing Double-Spending attack) Let the adversary $\mathcal{A}$ to execute $Create, Register$, $Withdraw$ and $Deposit$ protocols with the $Oracle$ to get the coins and transactions that can be redeemed by $\mathcal{A}$. Then $\mathcal{A}$ broadcasts two transactions that redeem a same transaction, so that the two transactions have the same ``head'' and, with the security of Bitcoin system, the probability that these two transactions appear on the $Ledger$ is negligible. Thus in this game, $\mathcal{A}$ can succeed with negligible probability.

(Anonymity) In the definition of anonymity, we divide it into three parts OtR, StR and StO. For OtR, we assume that the users communicate with other with a secure channel, so in the $Spend$ protocol between two pairs of honest users (U_1, U_2) and (U_3, U_4) , where U pays $coin_1$ to U_2, U_3 pays $coin_2$ to U_4 and $Val(coin_1) = Val(coin_2)$. Thus, what the adversary $\mathcal{A}$ can observe are $\text{View}_1 = \{(\text{PutMoney}^{U_1} = H(T^{U_1}), \pi_1, Val(coin_1), \sigma_1), \text{ClaimMoney}^{U_2} = (H(\text{PutMoney}^{U_1})$, $\pi_2, Val(coin_1), \sigma_2)\}$ and $\text{View}_2 = \{(\text{PutMoney}^{U_3} = H(T^{U_3}), \pi_3, Val(coin_2), \sigma_3)$, $\text{ClaimMoney}^{U_4} = (H(\text{PutMoney}^{U_3})$, $\pi_4, Val(coin_2), \sigma_4)\}$. Then $\mathcal{A}$ communicates with $OracleUSpend(coin_1)$ or $UDeposit(coin_2)$ and the view of $\mathcal{A}$ is $\text{View}_3 = \{(\text{PutMoney}^{O} = H(T^{O}), \pi_5, Val(coin_1), \sigma_5), \text{ClaimMoney}^{\mathcal{A}} = (H(\text{PutMoney}^{O}), \pi_6, Val(coin_1), \sigma_6)$ or $\text{View}_4 = \{(\text{PutMoney}^{O} = H(T^{O}), pk_{\mathcal{A}}, Val(coin_2), \sigma_7)\}$. Because $\text{View}_1 = \text{View}_2$, so $\mathcal{A}$ cannot link View_3 or View_4 to View_1 or View_2. As a result, in this game, $\mathcal{A}$ can succeed with negligible probability.

The same analysis method can be applied the to anonymity of StO and StR.

(Transferability, divisibility and fair exchange between users) The security analysis of transferability, divisibility and fair exchange can be reduced to the security of Bitcoin system. In this system, (1) the payee can use the coin further without depositing to the bank first by redeeming the transactions on the $Ledger$; (2) different coins have different denominations and each of them is transferred by an unique transactions so that user can spend arbitrary amount of coins; (3) the detailed form of bitcoin transaction allow us to define the conditions that the transactions can be redeemed. Thus, no user can cheat the others.

6. CONCLUSION

In this work, we develop a more practical E-Cash system that can be used to achieve unforgeability, preventing double-spending attacks, anonymity, transferability, divisibility, fixed size of coin during the transfer process and fair exchange between users. Our main contribution is to achieve the above security properties in one E-cash system. And the further work is to reduce complexity, e.g., in the Withdraw protocol, the bank signs every coin, so we assume that the amount of coins that the user can withdraw is limited and to achieve the real arbitrariness is the future works.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their valuable comments. This work is supported in part by the National Key R&D Program of China (No. 2020YFB1005801) and in part by the National Natural Science Foundation of China (No. 62172396).

REFERENCES

[1] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[J]. Consulted, 2009.

[2] Chaum D. Blind Signatures for Untraceable Payments[M]// Advances in Cryptology. Springer US, 1983:199-203.

[3] Baldimtsi F, Chase M, Fuchsbauer G, et al. Anonymous Transferable E-Cash[M]// Public-Key Cryptography -- PKC 2015. 2015:101-124.

[4] Okamoto T, Ohta K. Disposable Zero-Knowledge Authentications and Their Applications to Untraceable Electronic Cash[C]// on Advances in Cryptology. Springer-Verlag New York, Inc. 1989:481-496.

[5] Okamoto T, Ohta K. Universal Electronic Cash[C]// Advances in Cryptology - CRYPTO '91, International Cryptology Conference, Santa Barbara, California, Usa, August 11-15, 1991, Proceedings. DBLP, 1991:324-337.

[6] Chaum D, Pedersen T. Transferable cash grows in size[J].

[7] Tewari H, Hughes A. Fully Anonymous Transferable Ecash[M]// Public-Key Cryptography -- PKC 2015. Springer Berlin Heidelberg, 2016:101-124.

[8] Märtens P. Practical compact e-cash with arbitrary wallet size[J]. Cryptology ePrint Archive, 2015.

[9] Man H A, Susilo W, Mu Y. Practical Compact E-Cash[M]// Information Security and Privacy. Springer Berlin Heidelberg, 2007:431--445.

[10] Camenisch J, Hohenberger S, Lysyanskaya A. Compact E-Cash[J]. Eurocrypt, 2005, 3494:566-566.

[11] Canard S, Pointcheval D, Sanders O, et al. Divisible E-Cash Made Practical[M]// Public-Key Cryptography -- PKC 2015. Springer Berlin Heidelberg, 2015:77-100.

[12] Heilman E, Baldimtsi F, Goldberg S. Blindly Signed Contracts: Anonymous On-Blockchain and Off-Blockchain Bitcoin Transactions[M]// Financial Cryptography and Data Security. Springer Berlin Heidelberg, 2016.

[13] Christidis K, Devetsikiotis M. Blockchains and Smart Contracts for the Internet of Things[J]. IEEE Access, 2016, 4:2292-2303.

[14] Kumaresan R, Bentov I. Amortizing Secure Computation with Penalties[C]// ACM Sigsac Conference on Computer and Communications Security. ACM, 2016:418-429.

[15] Kumaresan R, Moran T, Bentov I. How to Use Bitcoin to Play Decentralized Poker[C]// ACM Sigsac Conference on Computer and Communications Security. ACM, 2015:195-206.

[16] Kumaresan R, Bentov I. How to Use Bitcoin to Incentivize Correct Computations[C]// ACM Sigsac Conference on Computer and Communications Security. ACM, 2014:30-41.

[17] Banasik W, Dziembowski S, Malinowski D. Efficient Zero-Knowledge Contingent Payments in Cryptocurrencies Without Scripts[M]// Computer Security – ESORICS 2016. 2016.

[18] Andrychowicz M, Dziembowski S, Malinowski D, et al. Secure multiparty computations on Bitcoin[C]// Security and Privacy. IEEE, 2014:443-458.

[19] Srivastava, Shweta, and V. Saraswat. "E-Cash Payment Protocols." International Journal on Computer Science Engineering 4.9(2012).

[20] Anand R S, Madhavan C E V. An Online, Transferable E-Cash Payment System[M]// Progress in Cryptology —INDOCRYPT 2000. 2000:93-103.

AUTHORS

Peifang Ni received the Ph.D. degree in the Institute of Information Engineering, Chinese Academy of Sciences, in 2020. She is currently a postdoctoral with the Trusted Computing and Information Assurance Laboratory, Institute of Software, Chinese Academy of Sciences. Her major research interests include applied cryptography, security protocol and blockchain consensus.

AUGMENTED EFFICIENT ZERO-KNOWLEDGE CONTINGENT PAYMENTS IN CRYPTOCURRENCIES WITHOUT SCRIPTS

Peifang Ni[1, 2]

[1]TCA Laboratory, Institute of Software,
Chinese Academy of Sciences, Beijing, China
[2]State Key Laboratory of Cryptology, Beijing, China

ABSTRACT

Zero-Knowledge Contingent Payment presents how Bitcoin contracts can provide a solution for the so-called fair exchange problem.Banasik, W. et al. first presented an efficient Zero-Knowledge Contingent Payment protocol for a large class of NP-relations, which is a protocol for selling witness. It obtains fairness in the following sense: if the seller aborts the protocol without broadcasting the final message then the buyer finally gets his payment back. However, we find that the seller in the protocol could refuse to broadcast the final signature of the transaction without any compensation for the buyer. As a result, the buyer cannot get the witness from the final signature of the transaction and has the payment for the witness locked until finishing the large computation for a secret signing key.

In this paper, we fix this problem by augmenting the efficient Zero-Knowledge Contingent Payment protocol. We present a new protocol where the seller needs to provide the deposit before the zero-knowledge proof of knowledge of the witness being sold. And then the buyer could obtain the seller's witness if the seller broadcasts the final signature of the transaction and gets the payment and his deposit. Otherwise, the buyer could get back the payment and obtain the seller's deposit. This new augmented protocol is constructed without any new assumptions.

KEYWORDS

fair exchange, Bitcoin, cryptocurrencies, zero-knowledge, without scripts.

1. INTRODUCTION

The concept of cryptocurrency emerged in the last few years and recently there has been a huge emphasis on constructing cryptocurrencies easy for circulation. The main valuable property of these cryptocurrencies is that their security does not need to rely on any single trusted third party. Bitcoin [1], the most prominent of the cryptocurrencies, is a decentralized payment system that is based on maintaining a public transaction ledger in a distributed manner. The list of transactions in this payment system is written on a public ledger, which is maintained jointly by the system users. With the public ledger, the system can implement an idea of the so-called smart-contracts. Consider the Zero-Knowledge Contingent Payment [2], which is a contract and shows that how Bitcoin contracts can provide a solution for the so-called fair exchange problem. With respect to [2], the Zero Knowledge Contingent Payment makes it possible to make payments using Bitcoin in a trustless manner where neither the payer or payee can cheat and that the payments are given

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 53-67, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121605

to the payee only in the case that some knowledge is disclosed by the payee. The execution of this contract is guaranteed by the rules of the underlying Bitcoin system.

To be more specific, it is executed between two parties that do not trust each other: the Seller and the Buyer. The Buyer is looking for some value $x \in \{0,1\}^*$ and the valuable conditions of x for him can be described as a function $f: \{0,1\}^* \rightarrow \{true, false\}$ (in a form of a polynomial-time computer program), such that finding x satisfying $f(x) = true$ is much harder than verifying that $f(x) = true$ holds. Hence than, the Buyer is willing to pay for x in the conditions that x is valuable for him. Imagine now that the Buyer is approached by a Seller through the internet, who is claiming that he knows x satisfying the valuable condition $f(x) = true$ and willing to sell x. Then the parties face the following problem: they need to finish the transaction on the internet in a trustless manner where neither the Seller or the Buyer can cheat.

The Zero-Knowledge Contingent Payment protocol described in [2] is accomplished using the combination of a hash-locked transaction and some non-standard scripts so that the data revealed in the hashlock release is the data in need. Recently, It has been implemented in [3] for selling a proof of a sudoku solution.

Banasik, W. et al. first created non-trivial efficient smart contracts using only the standard transactions in the public ledger [4]. Under the assumption of semantically secure Paillier encryption and symmetric encryption, secure commitment and time-locked commitment schemes, the strongly unforgeable ECDSA scheme, and zero-knowledge proof of knowledge, they constructed efficient Zero-Knowledge Contingent Payment protocol for a large class of NP-relations, which is a protocol for selling the witness. In the protocol, the buyer first prepares a transaction T_1 sending the funds from his public key to a public key shared by the buyer and the seller, and then they sign the transaction T_2, sending the funds from the shared public key to the seller's public key, using secret-shared signing keys in cooperation with the seller knowing the final signature of the transaction but the buyer not. After the seller giving a zero-knowledge proof of the witness being sold, the buyer broadcasts T_1 and the seller broadcasts T_2. And then the buyer can reverse the witness from the final signature. However, we find that the seller in the protocol could refuse to broadcast the final signature of the transaction without any compensation for the buyer. As a result, the buyer cannot get the witness from the final signature of the transaction and has the payment for the witness locked until finishing the large computation for a secret signing key.

1.1. Our Results

In this paper, we fix this problem by augmenting the efficient Zero-Knowledge Contingent Payment protocol. We present a new protocol where the seller needs to provide the deposit before the zero-knowledge proof of knowledge of the witness being sold. And then the buyer could obtain the seller's witness if the seller broadcasts the final signature of the transaction and gets the payment and his deposit. Otherwise, the buyer could get back the payment and obtain the seller's deposit. This new augmented protocol is constructed without any new assumptions. A high-level overview is as the following.

The ECDSA signature scheme is denoted by $(ECGen, ECSign, ECVer)$. A user in Bitcoin system is identified by his public key pk in the ECDSA signature scheme and each key pk is called an address. A simple transaction denoted by $[T]$ simply sends some Bitcoins x from address pk_0 to pk_1 . $[T]$ contains the following tuple $[T] \coloneqq (TXid(T'), value: x, from: pk_0, to: pk_1)$, where $TXid(T')$ denotes the identifier of transaction$[T']$ with value at least x that appeared earlier on the ledger and is redeemed by $[T]$.

A complete transaction denoted by T has a form $([T], ECSign_{sk_0}([T]))$. $TXid(T')$ is defined simply as a $SHA256$ hash of $([T], ECSign_{sk_0}([T]))$.

Under the assumption of semantically secure Paillier encryption and symmetric encryption, secure commitment and time-locked commitment schemes, the strongly unforgeable ECDSA scheme, and zero knowledge proof, our augmented efficient Zero-Knowledge Contingent Payment protocol consists of four stages.

In stage 1, the buyer and seller execute a key exchange protocol to generate two key pairs for the ECDSA signatures such that the secret keys are secret-shared between them. As a result, the buyer holds (PK, SK_0) , (PK', SK'_0) and seller holds $(PK, SK_1), (PK', SK'_1)$.

In stage 2, the seller prepares transaction T_3 sending the funds p from PK_S to PK' and T'_3sending the funds p from PK' to PK_S, and then sends the hash value of T_3 to the buyer. The buyer prepares transaction T_1 sending the funds v from PK_B to PK', T'_1 sending the funds $v + p$ from PK' to PK, and T_2 sending the funds $v + p$ from PK to the seller's public key PK_S. They execute the unique signature generation protocol to generate the signature of $[T'_1]$ for the buyer and the signatures of T_2 and $[T'_3]$ for the seller. And then the seller broadcasts T_3. If T_3 is not corresponding to the hash value of sent before, the buyer aborts. Otherwise, the parties execute the following stages.

In stage 3, the seller proves the knowledge of the witness being sold by executing a zero-knowledge proof of knowledge protocol with the buyer in the cut-and-choose technique. In the proof, the seller uses the signature of $[T_2]$ to compute the secret key for encryption of a set of challenges and responses with the witness, and then it sends the commitments of the encryption to the buyer. After receiving the subset of the challenges from the buyer, the seller opens the commitments asked to open. The proof is valid if the output of the buyer's verification is true. Otherwise, the buyer aborts.

In stage 4, if the buyer broadcasts the valid T_1 and T'_1, and the seller broadcasts the valid T_2, the buyer could reverse the witness from the signature of T_2. If the buyer refuses to broadcast T_1 and T'_1 or he broadcasts illegal T_1 and T'_1, then the seller could broadcast T'_3 to get the deposit back. If the buyer broadcasts the valid T_1 and T'_1, but the seller refuses to broadcast T_2, the buyer could finally obtain his own funds together with the seller's deposit after finishing a large computation for a secret signing key.

Therefore, it holds that:

assume the existence of semantically secure Paillier encryption and symmetric encryption, secure commitment and time-locked commitment schemes, the strongly unforgeable ECDSA scheme, and zero knowledge proof of knowledge. Then, there exists a secure efficient Zero-Knowledge Contingent Payment protocol in cryptocurrencies without scripts, which obtains fairness in the following sense: if the seller aborts protocol without broadcasting the final message then the buyer finally gets its payment back and gets an extra financial compensation from the seller.

1.2. Related Works

Relevant to our work are the works on smart contracts that provide solutions for fair protocols in the cryptocurrency systems. Bentov, I. and Kumaresan, R. studied secure computations in the following model of fairness: a malicious user who aborts protocol after receiving the output is forced to pay a mutually predefined monetary penalty [5]. They then showed how to use Bitcoin system to achieve secure computations with the above defined fairness in two-party setting as well as the multiparty setting (with a honest majority) by simulating with a new ideal functionalities as they proposed.

Andrychowicz, M. et al. showed how to obtain fair two-party secure computation protocol via the Bitcoin system with the fairness in the following sense: if one party aborts the protocol after learning the output but the other one not, then the other party gets a financial compensation from the aborted one [6]. They constructed the protocol with the two-party protocol of Goldreich and Vainish [7] additionally secured against an active adversary with zero-knowledge proofs. And they presented one possible application of the protocol to the fair contract signing: each party is forced to either complete the protocol or pay a fine to the other party.

1.3. Outline

In Section 2, we define the notations and definitions that are used through the paper. In Section 3, we describe the subprotocols used in the Banasik's efficient Zero-Knowledge Contingent Payment and give the new subprotocols used in our protocol. In Section 4, we present our augmented efficient Zero-Knowledge Contingent Payment protocol. And Section 5 is conclusion.

2. PRELIMINARIES

2.1. Notations

We use n to denote the security parameter. We use $[k]$ for any $k \in N$ to denote the set $\{1, \dots, k\}$. For any probabilistic algorithm $A(\cdot)$, $A(x)$ is the result of executing A with input x and uniformly chosen randomness. We use $y = A(x)$ to denote that y is set to $A(x)$. For a set S, we use $y \in S$ (or $y \leftarrow S$) to denote that y is chosen from S. For any language L) and instance$x \in L$), we use $\{\mathcal{R}_L\}(x)$) to denote the set of witnesses for $x \in L$. A function $\mu: N \to R$ is negligible (in x) if for every positive integer c there exists an integer N_c such that for all x>N_c we have that $|\mu(x)| < \frac{1}{x_c}$.

2.2. Elliptic Curve Digital Signature Algorithm (ECDSA)

We recall the definition of digital signature algorithm [8, 4, 9].

Definition 1 (Digital Signature Scheme): A digital signature scheme consists of a triple of probabilistic polynomial-time algorithms $(Gen, Sign, Ver)$ satisfying the following conditions:

- Key-generator algorithm Gen: On input 1^n, Gen outputs a pair of keys(sk, vk);
- Signing algorithm $Sign$: On input secret key sk and a message α, $Sign$ outputs signatureσ;
- (Deterministic) Verifying algorithm Ver: On input public key vk, message α, and signature, σ , Ver returns $Output \in \{\text{accept}, \text{reject}\}$;
- For every pair (sk, vk) in the range of $Gen(1^n)$ and for every $\alpha \in \{0,1\}^*$, the signing and verification algorithms $Sign$ and Ver satisfy
$$\Pr\left[Ver\left(vk, \alpha, Sign(sk, \alpha)\right) = 1\right] = 1$$

where the probability is taken over the internal coin tosses of algorithms $Sign$ and Ver.

Definition 2 (Elliptic Curve Digital Signature Algorithm) The ECDSA signature scheme $(ECGen, ECSign, ECVer)$ is a variant of digital signature scheme algorithm, in which the algorithms $ECGen, ECSign, ECVer$ are defined as the followings.

- Key-generator algorithm $ECGen$: On input 1^n, $ECGen$ chooses an elliptic curve group $(G, O, g, +)$ over a prime field Z_p, where O is the neutral element, g is the generator of G, and the order of G is a prime number such that $\lceil log_2|G|\rceil = n$. And then $ECGen$ samples a random $d \in_R Z_{|G|}$, and computes $D \coloneqq d \cdot g$. The generated secret key is $(d, (G, O, g, +))$, and the public key is $(D, (G, O, g, +))$;
- Signing algorithm $ECSign$: Let H be a hash function and $f: G \to Z_{|G|}$ be a reduction function that we will define in a moment. On input secret key and messageα, $ECSign$ first chooses a random $k \in_R Z_{|G|}$, and then computes $r = f(k \cdot g)$ and $s = k^{-1}(H(\alpha) + d \cdot r) mod|G|$. If $r = 0$ or $s = 0$ then the algorithm aborts. Otherwise, the $ECSign$ outputs (r, s);
- (Deterministic) Verifying algorithm $ECVer$: On input public key, message α and signature (r, s), $ECVer$ first checks if r and s are non zero elements of $Z_{|G|}$ and then verifies if $r = f(H(\alpha) \cdot s^{-1} \cdot g + r \cdot s^{-1} \cdot D)$. If this holds, then $ECVer$ outputs ok, and otherwise it outputs$\perp$.

Recall that every element of G has the form $(x, y) \in Z_p^2$. The reduction function $f: G \to Z_{|G|}$ is defined as: on input (x, y) ignores y and produces as output $f(x, y) = xmod\ \ |G|$.

Security of the ECDSA signature scheme. The ECDSA signature scheme satisfies existentially strong unforgeability under a chosen message attack. We consider the following game played by a polynomial time adversary$A$. First, a key pair is sampled as $(vk, sk) \coloneqq ECGen(1^n)$ and the adversary A is given the public verification key vk. Then, A chooses a sequence of messages $\alpha_1, \dots, \alpha_k$ and learns each corresponding signature $\sigma_i \coloneqq ECSign_{sk}(\alpha)$. He does it in an adaptive way that he chooses each α_i after learning $\sigma_1, \dots, \sigma_{i-1}$. Finally, A outputs a pair $(\alpha_{k+1}, \sigma_{k+1})$. We say that A mauls a signature if the output $(\alpha_{k+1}, \sigma_{k+1})$ satisfies that $ECVer(\alpha_{k+1}, \sigma_{k+1}) = ok$ and σ_{k+1} has not been sent to A as one of the signatures in $\sigma_1, \dots, \sigma_k$. The ECDSA signature scheme is existentially strongly unforgeable under a chosen message attack if every polynomial-time adversary can maul a signature with at most negligible probability.

2.3. Homomorphic Encryption Schemes

We recall the definitions of encryption schemes in [4, 10, 11] and present the following definition.

Definition 3 (Public Encryption Scheme) A public encryption scheme consists of a triple of probabilistic polynomial-time algorithms $(EncGen, Enc, Dec)$:

- Key-generator algorithm $EncGen$: On input 1^n, $EncGen$ outputs a pair of keys (pk, sk);
- Encryption algorithm Enc: On input public key pk and message $\alpha \in \{0,1\}^*$, Enc outputs ciphertext $c = Enc_{pk}(\alpha) \in \{0,1\}^*$;
- Decryption algorithm Dec: On input private key sk, ciphertext c, Dec outputs $\alpha' = Dec_{sk}(c)$;
- For every pair (pk, sk) in the range of $EncGen(1^n)$ and for every $\alpha \in \{0,1\}^*$, the encryption and decryption algorithms Enc and Dec satisfy

$$\Pr\left[Dec_{sk}\left(Enc_{pk}(\alpha)\right) = \alpha\right] = 1$$

where the probability is taken over the internal coin tosses of algorithms Enc and Dec.

Security. To define security of an encryption scheme $(EncGen, Enc, Dec)$ consider a polynomial time adversary A produces a pair of messages (m_0, m_1) with access to oracle $Enc_{pk}(\cdot)$ and $Dec_{sk}(\cdot)$, where $(pk, sk) = EncGen(1^n)$. He then receives a ciphertext $c = Enc_{pk}(m_b)$ for a random $b \in \{0,1\}$, and produces $b' \in \{0,1\}$ without asking the oracle $Dec_{sk}(\cdot)$ to decrypt c. We say that A wins if $b' = b$. We say that the encryption scheme $(EncGen, Enc, Dec)$ is semantically secure if for every polynomial time adversary A the probability that he wins is at most$\frac{1}{2} + \mu(n)$, where μ is some negligible function (in other words: $Enc_{pk}(m_0)$ and $Enc_{pk}(m_1)$ are computationally indistinguishable).

The definition of a symmetric-key encryption scheme is the same as that of a public encryption scheme except that there is only one secret key k, which is usually sampled uniformly at random from some space $\mathcal{K}$ (that depends on 1^n). The adversary does not learn the key, but can try to get some partial information about k by choosing messages $m_1, \dots, m_n$ and learning the corresponding ciphertexts $Enc_{pk}(m_1), \dots, Enc_{pk}(m_n)$ in an adaptive way.

A public-key encryption scheme $(EncGen, Enc, Dec)$ is additively homomorphic if the set of valid messages for the public key pk is an additive group $(H_{pk}, +)$, where the key pair (pk, sk) is generated by $EncGen$. Moreover, the homomorphic algorithm we require is defined by an operation $\otimes: \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^* \{\perp\}$, such that for every valid key pair (pk, sk) and every pair of messages $(m_0, m_1) \in H_{pk}$ we have that $Dec_{sk}\left(Enc_{pk}(m_0) \otimes Enc_{pk}(m_1)\right) = m_0 + m_1$.

2.4. Time-lock Commitment Schemes

We recall the definitions of time-lock commitment schemes [4, 12, 13]. A commitment scheme $(Commit, Open)$ executed between two parties (a committer and a receiver) consists of two phases, the commit phase and the open phase.

- Commit Phase. The committer takes as input a message $\alpha \in \{0,1\}^*$ and chooses a random $r \in \{0,1\}^*$. The committer then computes $c := Commit(\alpha)$ with randomness r and sends it to the receiver;
- Open Phase. The committer reveals the message α committed and the randomness r used in the commit phase. And the receiver verifies $Open(c, \alpha, r)$.

Security. A commitment scheme is secure if it is binding and hiding. The hiding property means that for every $\alpha_0, \alpha_1 \in \{0,1\}^*$, $Commit(\alpha_0)$ and $Commit(\alpha_1)$ are computationally indistinguishable. The binding property means that for any PPT committer, he can reveal another message $\alpha' \neq \alpha$ such that $Open(c, \alpha, r) = \alpha'$ with negligible probability.

A commitment scheme $(Commit, Open)$ is a time-lock commitment scheme if the receiver can open the commitment by himself with a significant computational effort. Every time-lock commitment comes with two parameters, τ_0 and τ_1 (with $\tau_0 \leq \tau_1$), and is called (τ_0, τ_1)-secure, where τ_0 denotes the time that everybody, including very powerful adversaries, needs to force open the commitment, and τ_1 denotes the time needed by the honest users to force open the commitment.

2.5. Zero-knowledge Proofs of Knowledge

We recall the definitions of zero-knowledge protocols [4, 14, 15, 16] and give a simple introduction.

Definition 4 (Interactive Proof System) A pair of interactive Turing machines $\langle P, V\rangle$ is called an interactive proof system for any language L if machine V is polynomial-time and the following two conditions hold:

- Completeness: There exists a negligible function c such that for every $x \in L$,

$$\Pr[\langle P, V\rangle(x) = 1] > 1 - c(|x|)$$

- Soundness: There exists a negligible function s such that for every $x \notin L$ and every interactive machine B, it holds that

$$\Pr[\langle P, V\rangle(x) = 1] < s(|x|)$$

$c(\cdot)$ is called the completeness error and $s(\cdot)$ is the soundness error.

Zero-knowledge protocols are interactive proof systems with zero-knowledge property, which means that the prover knows a witness ω for some instance $x \in L$ and convinces the verifier that $x \in L$ without providing the verifier with any additional information beyond the fact that $x \in L$.

A zero-knowledge protocol is called a zero-knowledge proof of knowledge if $L \in NP$ and for every prover P^* there exists a polynomial-time knowledge extractor, that can output a witness ω for $x \in L$ by interacting with P^*. We follow the requirement in [4], suppose that the last two messages in the zero-knowledge proof of knowledge protocol are: a challenge c sent by the verifier to prover, and the prover's response r corresponding to the challenge c. The knowledge extractor extracts witness ω after being given transcripts of two accepting executions that are identical except that the challenges are different (and the responses may also be different).

2.6. Bitcoin Transaction Syntax

A complete transaction denoted by T has a form $([T], ECSign_{sk_0}([T]))$, where $[T] \coloneqq (TXid(T'), value{:}\, x, from{:}\, pk_0, to{:}\, pk_1)$. Another standard type of the transaction is called multisig transaction [4], where $[T]$ has a form $(TXid(T'), value{:}\, x, from{:}\, pk_0, to\ \ k - out - ofm{:}\ \ pk_1, \dots, pk_m)$ and it is signed by sk_0. It can be redeemed by a transaction T'' with the form $([T''], \sigma_{i_1}, \dots, \sigma_{i_k})$, $1 \le i_1 < \cdots < i_k \le m$ and for every $1 \le j \le k$ it holds that $ECVer_{pk_{i_j}}\left([T''], \sigma_{i_j}\right) = ok$.

3. THE SUBPROTOCOLS IN THE ZERO-KNOWLEDGE CONTINGENT PAYMENT PROTOCOL

In this section, we describe the subprotocols used in Banasik's efficient Zero-Knowledge Contingent Payment system and give the new formalization used in our protocol.

3.1. The Two-party ECDSA Key Generation Protocol

The first ingredient is a protocol called $SharedKGen$, in which two parties, the Seller and Buyer, generate a key pair (public key, private key) for ECDSA, in such a way that the secret key is secret-shared between the Seller and Buyer. This protocol is still used in our final protocol.

To be more precise, fix an elliptic curve $(G, Og, +)$ constructed over a prime field Z_p and let $Com = (Commit, Open)$ be a secure (computational hiding) commitment scheme, both parties take as input a security parameter 1^n, the overview of $SharedKGen$ is as following:

- Seller: samples $d_S \leftarrow Z^*_{|G|}$, computes $D_S \coloneqq d_S \cdot g$ and $c \coloneqq Commit(D_S)$, and sends c to the Buyer
- Buyer: samples $d_B \leftarrow Z^*_{|G|}$, computes $D_B \coloneqq d_B \cdot g$ and sends D_B to the Seller;
- Seller: sends $Open(D_S)$ to the Buyer;
- The Seller and Buyer compute the ECDSA public key separately: $pk \coloneqq d_S \cdot D_B = d_B \cdot D_S$ and abort if pk =0. And the secret key $sk := d_S \cdot d_B$ is secret-shared between the two parties.

3.2. The Unique Signature Generation Protocol

We will present two formalizations of the unique signature generation protocol, USG_1 and USG_2. USG_1 is a little different from the USG protocol used in Banasik's efficient Zero-Knowledge Contingent Payment system [4].

The USG protocol uses $KSignGen$ as a subroutine, in which the parties jointly sign a message z using the secret-shared signing key generated by $SharedKGen$. Recall that G is an elliptic curve group for ECDSA signature scheme $ECGen, ECSign, ECVer$, $TLCom = (TLCommit, TLForceOpen)$ is a time-locked commitment scheme, and $(AddHomGen, \mathrm{AddHomEnc}, \mathrm{AddHomDec})$ is a Paillier encryption scheme which is additively homomorphic over $Z_{n'}$, where $n' > 2 \cdot |G|^4$. Let F be a one-way function. Protocol KSignGen is described as follows.

- Seller and Buyer: They jointly create signing randomness K using the same way in $SharedKGen$. The Seller holds $K_S \in Z^*_{|G|}$ and $K_S \coloneqq k_s \cdot g$ and the Buyer holds $K_B \in Z^*_{|G|}$ and $K_B \coloneqq k_B \cdot g$. Finally the parties both know the signing randomness $K \coloneqq k_s \cdot k_B \cdot g$, parse K as (x, y), compute $r \coloneqq xmod\ \ |G|$, and abort if$r = 0$;
- Seller: He creates a new key pair $(pk_{AH}, sk_{AH}) \coloneqq \mathrm{AddHomGen}(1^n)$ in the Paillier encryption scheme, computes the ciphertext $c_S \coloneqq \mathrm{AddHomEnc}_{pk_{AH}}(d_S)$, and sends the public key pk_{AH} and the ciphertext c_S;
- Buyer: On receiving pk_{AH} and c_S from the Seller, he computes $c_0 \coloneqq k_B^{-1} \cdot H(z) mod|G|$, $c_1 \coloneqq \mathrm{AddHomEnc}_{pk_{AH}}(c_0)$, $t \coloneqq k_B^{-1} \cdot r \cdot d_B mod\ \ |G|$, $c_2 \coloneqq c_1 \otimes (c_S)^t$, samples $u \leftarrow \{1, \ldots, |G|^2\}$, computes $c_B \coloneqq c_2 \otimes \mathrm{AddHomEnc}_{pk_{AH}}(u \cdot |G|)$, and sends the ciphertext c_B;
- Seller: On receiving c_B, he computes $S_0 \coloneqq \mathrm{AddHomDec}_{sk_{AH}}(c_B)$, $s \coloneqq k_S^{-1} s_0\ mod\ \ |G|$ and aborts if s=0. The final signature of z is $\sigma \coloneqq (r, s)$ if $\mathrm{ECVer}_{pk}(z, \sigma) = ok$ and otherwise the Seller aborts. At the end he commits to $S = F(\sigma)$, creates a time-lock commitment to d_S, and sends $\Gamma_i \coloneqq Commit(S)$ and $\Phi_i \coloneqq \mathrm{TLCommit}(d_S)$;

We stress that when the two parties execute $KSignGen$, the one who is supposed to obtain the final signature at the end of $KSignGen$ plays the role of ``Seller".

The USG protocol is executed after the parties generate a key pairs $(sk^1, pk^1), \ldots, (sk^a, pk^a)$ using the SharedKGen protocol.

- Buyer: He chooses a random subset $J \subseteq \{1, \ldots, a\}$, such that $|J| = a - b$. Let $j_1, \ldots, j_b$ denote the set $\{1, \ldots, a\} \backslash J$. And then he chooses message z to be signed and sends it to the Seller；
- Seller and Buyer: For $i = 1$, the parties execute the $KSignGen(1^n)$ protocol with $pk_i \coloneqq d_S^i \cdot d_B^i \cdot g$. And at the end of each execution, the Seller sends the commitment $\Gamma_i \coloneqq \text{Commit}(S^i)$, where $S^i = F(\sigma_i)$, and the time-lock commitment $\Phi_i \coloneqq \text{TLCommit}(d_S^i)$；
- Buyer: He sends J to Seller；
- Seller: For each $j \in J$, the Seller opens the commitments to S^j and d_S^j, and sends σ^j, k_S^j and sk_{AH}^j to Buyer；
- Buyer: He aborts if any of the commitments did not open correctly. Otherwise, for each $j \in J$, he verifies if the following holds:
  - $\text{ECVer}_{pk^j}(z, \sigma^j) = ok$;
  - $F(\sigma^j) = S^j$;
  - $d_S^j \cdot d_B^j \cdot g = pk^j$;
  - $\text{AddHomDec}_{sk_{AH}}^j(c_S^j) = d_S^j$.

  If the verification fails then the Buyer aborts. Otherwise, the unique signature is the set $\{\sigma^{j_i}\}$, where $i \in \{1, \ldots, b\}$.

In this paper, the USG_1 protocol is the same as the USG protocol except that USG_1 outputs the set $\{\sigma^i\}$, where $i \in \{1, \ldots, a\}$, as the unique signature. For the technique used in the final protocol, we assume that each $S^i$from the USG_1 protocol is divided into $2n$ parts $S^{i,1}, \ldots, S^{i,2n}$ each of size n. In addition, we assume that each part $S^{i,j}$ is committed separately.

The USG_2 protocol is the same as the $\text{KSignGen}(1^n)$ protocol except that there is no further computation after obtaining the signature $\sigma \coloneqq (r, s)$, of which the verification result is ok.

4. The Augmented Zero-Knowledge Contingent Payment Protocol

In this section we show how to use the subprotocols to construct the augmented two-party Zero-Knowledge Contingent Payment Protocol protocol with fairness in the following sense: if the malicious seller aborts protocol without broadcasting the final message then the buyer finally gets his payment back and an extra financial compensation from the seller.

4.1. Security Definition

In our protocol, the Seller sells to the Buyer a value x such that $f(x) = true$ (for some public function $f: \{0,1\}^* \to \{true, false\}$. We assume that the price of x is v Bitcoins and the deposit of the Seller for this transaction is p Bitcoins. Hence, before an execution of the protocol starts, there is some unspent transaction T_0 on the blockchain that sends v Bitcoins to pk_B and some unspent transaction T'_0 on the blockchain that sends p Bitcoins to pk_S. The parties initially share the following common input: security parameter 1^n, price v for the secret x, deposit value$p$ for the computation of a time-lock committed message, parameters $a, b \in N$ such that $a > b$, an elliptic curve group $G, O, g, +$ for an ECDSA signature scheme, such that $\lceil log_2|G| \rceil =$

n, and parameters (τ_0, τ_1) for the time-lock commitment. We say that the final protocol is $\varepsilon - secure$ if, except with probability $\varepsilon + \mu(n)$, the following properties hold:

- if the honest Buyer loses his funds then he learns x'such that $f(x') = true$;
- if the honest Seller does not get Buyer's funds then the Buyer learns no information about x;
- if an honest Buyer is forced to open a time-lock commitment then he finally does not lose his funds and obtains a financial compensation from the Seller.

4.2. Instantiations and Assumptions

Instantiations. F is a one-way function and we use a standard symmetric encryption scheme $\mathrm{EncGen, Enc, Dec}$ and the additively-homomorphic public key encryption scheme $(\mathrm{AddHomGen,\ AddHomEnc, AddHomDec})$ introduced by Pascal Paillier [11]. The elements on which we perform the addition operations are the exponents in the elliptic curve group of the ECDSA scheme. Hence the Paillier encryption scheme is homomorphic over $Z_{n'}$, where $\mathrm{n}' > 2 \cdot |\mathrm{G}|^4$, and $\lceil log_2 |G| \rceil = n$.

We use a standard commitment scheme $\mathrm{Com} = (\mathrm{Commit, Open})$ that is based on hash function and secure in the random oracle model [4]. Let H be a hash function and a commitment to message x is defined as Commit (x):=H(x||r), where $r \in \{0,1\}^*$. And Open is to reveal(x, r). The binding property follows from the collision-resistance of H because that a commitment that can be open in two different ways would form a collision for H. And the hiding property follows from the fact that H(x||r) does not reveal any information about x. We use the classic (τ_0, τ_1)-secure timed commitments TLCom $=$ $(\mathrm{TLCommit, TLForceOpen})$ in [16] and assume that $\tau_1 = 10 \cdot \tau_0$ as in [4].

We consider the two-party protocol, executed between a Buyer and a Seller, in which the Seller sells to the Buyer x such that $f(x) = true$, in the active security settings.The parties are connected by a secure channel, which can be easily obtained using the standard cryptographic techniques. One user in Bitcoin is identified by his public key in the ECDSA signature scheme, which helps to establish the secure channel between each other. Without loss of generality, we set (PK_B, SK_B) and (PK_S, SK_S) to be the ECDSA key pairs of the Buyer and Seller respectively.

Assumptions. We assume that Paillier encryption and symmetric encryption are semantically secure, Com and TLCom are secure commitment schemes, and the ECDSA scheme is strongly unforgeable. Hence, the subprotocols in Section 3 are secure.

We keep the form of the assumption of zero-knowledge proof of knowledge protocol in [4]. We also assume that the public function f has a zero-knowledge proof of knowledge protocol, denoted by $\mathcal{F}$, in which the Seller can prove the knowledge of x such that $f(x) = true$. Protocol $\mathcal{F}$ consists of two phases: the $Setup_{\mathcal{F}}$ phase and the $Challenge_{\mathcal{F}}$ phase. After executing the$Setup_{\mathcal{F}}$ phase, the views of the Seller and Buyer are denoted by $S_{\mathcal{F}}$ and $B_{\mathcal{F}}$ respectively.

In the $Challenge_{\mathcal{F}}$ phase, the Buyer generates challenge message as $c_{\mathcal{F}} = \mathrm{GenChallenge}_{\mathcal{F}}$ $(B_{\mathcal{F}})$ and sends $c_{\mathcal{F}}$ to Seller. Then the Seller calculates the corresponding response $r_{\mathcal{F}}$ $= \mathrm{GenResponse}_{\mathcal{F}}(x, S_{\mathcal{F}}, c_{\mathcal{F}})$ and sends $r_{\mathcal{F}}$ to the Buyer. At the end, the Buyer decides to accept or reject the proof according to the output of function $=$ $\mathrm{VerifyResponse}_{\mathcal{F}}(B_{\mathcal{F}}, c_{\mathcal{F}}, r_{\mathcal{F}}) \in \{\mathrm{true, false}\}$.

Protocol $\mathcal{F}$ is a proof of knowledge since we require that there is a knowledge extractor $\text{Extract}_{\mathcal{F}}$ such that $\text{Extract}_{\mathcal{F}}(B_{\mathcal{F}}, c_{\mathcal{F}}^1, r_{\mathcal{F}}^1, c_{\mathcal{F}}^2, r_{\mathcal{F}}^2)$ and $f(x') = true$ if only $=$ $\text{VerifyResponse}_{\mathcal{F}}\left(B_{\mathcal{F}}, c_{\mathcal{F}}^i, r_{\mathcal{F}}^i\right) = true$ for $i = 1{,}2$ and $c_{\mathcal{F}}^1 \neq c_{\mathcal{F}}^2$. It means that a witness for f can be extracted from the correct answers corresponding to two different challenges. We also assume that the challenge $c_{\mathcal{F}}$ is sampled uniformly from the set $X_{A_{\mathcal{F}}} = \{0{,}1\}$ in Seller's view.

4.3. The Protocol

Our protocol consists of four stages and the final protocol called SellWitness is depicted on Fig.1.

- **Stage 1.** Using the two-party ECDSA key generation protocol SharedKGen, the Buyer and Seller jointly generate $a + 1$ key pairs $(sk^1, \text{pk}^1), \ldots, (sk^{a+1}, \text{pk}^{a+1})$, where $\text{pk}^i := d_S^i \cdot d_B^i \cdot g$ and $sk^i \coloneqq d_S^i \cdot d_B^i$. As a result, the Buyer holds $(PK, SK_0) \coloneqq ((\text{pk}^1, \ldots, \text{pk}^a),$
- $(d_B^1, \ldots, d_B^a))$, $(PK', SK'_0) \coloneqq (\text{pk}^{a+1}, d_B^{a+1}))$ and the Seller holds $(PK, SK_1) \coloneqq ((\text{pk}^1, \ldots, \text{pk}^a), (d_S^1, \ldots, d_S^a))$, $(PK', SK'_1) \coloneqq (\text{pk}^{a+1}, d_S^{a+1}))$;
- **Stage 2.** The parties respectively produce the messages to be signed. If one prepares a transaction T, then the message to be signed is $[T]$. The procedure is called GenMsg_T. The Seller prepares transaction T_3 sending the funds p from PK_S to PK' and T'_3 sending the funds p fromPK' to PK_S, and then sends the hash value ofT_3 to the Buyer. And then the Buyer prepares transaction T_1 sending the funds v from PK_B to PK', T'_1 sending the funds $v + p$ from PK' to a-out-of-$a + b - 1 PK, PK_B, \ldots, PK_B$ (consists of $b - 1 PK_B$), and T_2 sending the funds $v + p$ from PK to PK_S. They execute the unique signature generation protocol USG_2 using (PK', SK'_0) and (PK', SK'_1) to generate the signature of$[T'_1]$for the Buyer and the signature of $[T'_3]$ for the Seller. And then they execute the USG_1 using (PK, SK_0) and (PK, SK_1) to generate the signatures of $[T_2]$ for the Seller, in which the Seller will commit to SK_1 using TLCom and send to the Buyer. And then the Seller broadcasts T_3. If T_3 is not corresponding to the hash value of sent before, then the Buyer aborts; otherwise, the parties execute the following stages;
- **Stage 3.** The Seller proves the knowledge of x such that $f(x) = true$ by executing a zero-knowledge proof of knowledge protocol $\mathcal{F}$ with the Buyer in the cut-and-choose technique;
- **Stage 4.** The Buyer either extracts x such that $f(x) = true$ or takes his funds back and obtains the Seller's deposit.

Theorem. Suppose Paillier encryption and symmetric encryption are semantically secure, Com and TLCom are secure commitment schemes, and the ECDSA scheme used in the construction of the USG_1 and USG_2 is strongly unforgeability. Additionally, there is a zero knowledge proof $\mathcal{F}$ of knowledge of x s.t. $f(x) = true$ of the required form. Then the SellWitness (Fig.1) is ε-secure for $\varepsilon = (b/a)^b$.

Proof. Recall that SellWitness is ε -secure if, except with probability $\varepsilon + \mu(n)$, the following properties hold: (1) if an honest Buyer loses funds then he learns x' such that $f(x') = true$; (2) if an honest Seller does not get Buyer's funds then the Buyer learns no information about x; (3) if an honest Buyer is forced to open a time-lock commitment, then he finally does not lose his funds and obtains a financial compensation from the Seller. Hence then, we show the security analysis of SellWitness in two cases: (i) the Buyer is honest and the Seller is mailicious, and (ii) the Seller is honest and, while the Buyer is malicious.

At the beginning of this proof, we stress that the security of the unique signature generation protocol USG_1 follows that of the USG protocol used in Banasik's efficient Zero-Knowledge

Contingent Payment system [4], and the security of the USG_2 protocol follows that of the ECDSA.

For case (i), the honest Buyer loses his funds only if he broadcasts transactions T_1 and T'_1 is redeemed by the Seller or it just locks the Buyer's funds forever. Since with probability$\geq 1 - (b/a)^b - \mu_0(n)$ (for a negligible $\mu_0(n)$) at least one of the b chosen executions, consist of execution i for $i \in \{1, \dots, a\} \backslash j$, of the KSignGen procedure was completed honestly by the Seller (guaranteed by the security of USG_1 protocol). While if the Seller does not redeem T'_1, the honest Buyer will be able to redeem transaction T'_1 and get back his funds together with the Seller's deposit after that the Buyer force-opens one time-locked puzzle Φ_i in a honestly completed execution i of KSignGen, where $i \in \{1, \dots, a\} \backslash j$. Hence then, the property (3) follows and the Buyer's funds cannot be locked forever except with probability $(b/a)^b - \mu_0(n)$.

Assume that the Seller redeems transaction T'_1 and he can redeem T'_1 only via transaction T_2. Then the Buyer can use signatures $\hat{\sigma}_i$ to calculate secrets $S^{i,j}$. Then he decrypts all the values $\sigma^{i,j}$ to get all the challenges and responses $c_k^{i,j}$,$r_k^{i,j}$. At the end using any pair of responses he can calculate $x' = \text{Extract}_{\mathcal{F}}(B_{\mathcal{F}}^{i,j}, c_1^{i,j}, r_1^{i,j}, c_2^{i,j}, r_2^{i,j})$ s.t. $f(x') = true$ with probability $\geq 1 - \mu_1(n)$, where $\mu_1(n)$ is negligible (see the proof of Lemma 2 in [4]). Hence then, the property (1) follows.

For case (ii), if an honest Seller does not get the Buyer's funds, the property (2) is guaranteed by the zero-knowledge property of the zero-knowledge proof of knowledge protocol.

Common Input: function f: $\{0,1\}^* \rightarrow \{\text{true}, \text{false}\}$.
Private Input: Seller holds value x such that $f(x) = true$.

- **Stage 1:**
  - The parties execute the SharedKGen protocol $\text{a} + 1$ times using the provided parameters. As a result, the Buyer holds $(PK, SK_0), (PK', SK'_0)$ and the Seller holds$(PK, SK_1), (PK', SK'_1)$ as defined previously.
- **Stage 2:**
  - The Seller generates transactions T_3 and T'_3, and then sends the hash value of T_3 to the Buyer;
  - The Buyer generates transactions T_1,T'_1 and T_2;
  - The Seller and buyer executeUSG_1 and USG_2 with the keys generated in Stage 1 to generate the corresponding signatures of $[T'_1]$, $[T'_3]$ and $[T_2]$. And the Buyer holds complete transaction T'_1 and Seller holds complete transactions T'_3 and T_2, and the unique signature $\{\sigma^i\}$ $(i \in \{1, \dots, a\})$ of $[T_2]$;
  - The Seller sends commitment $\Gamma_i := \text{Commit}(S^i)$ of $S^i = \mathcal{F}(\sigma^i)$ and the time-lock commitment $\Phi_i := \text{TLCommit}(d_S^i)$. Each S^i is divided into 2n parts $S^{i,1}, \dots, S^{i,2n}$ with size n and each part $S^{i,j}$ is committed separately. Let $\{S^i\}$, where $i \in \{1, \dots, a\} \backslash J$, denote the unopened set of $\Gamma_i := \text{Commit}(S^i)$. And the Seller also broadcasts T_3. If T_3 is not corresponding to the hash value sent before, the Buyer aborts. Otherwise, the parties execute the following stages.
- **Stage 3:** for $i \in \{1, \dots, a\} \backslash J$,
  - For $j \in \{1, \dots, 2n\}$, the Seller and Buyer execute the $\text{Setup}_{\mathcal{F}}^{i,j}$ phase and learn the views $\text{S}_{\mathcal{F}}^{i,j}$ and $\text{B}_{\mathcal{F}}^{i,j}$ respectively;
  - For $j \in \{1, \dots, 2n\}$, the Seller calculates two challenges $c_1^{i,j}$ and $c_2^{i,j}$ in random order, which will be chosen by the Buyer later to calculate the responses $r_k^{i,j} = \text{GenRespose}_{\mathcal{F}}(x, B_{\mathcal{F}^{i,j}}, c_k^{i,j})$ for $k = 1.2$;
  - For $j \in \{1, \dots, 2n\}$,, the Seller uses $\text{S}_{\mathcal{F}}^{i,j}$ to encrypt $\Gamma_k^{i,j} = \text{Enc}_{S^{i,j}}(c_k^{i,j}, r_k^{i,j})$ for $k = 1,2$ And he commits to $\gamma_k^{i,j}$for $k = 1.2$;
  - The Buyer chooses random subset $\mathcal{J}^i \subseteq \{1, \dots, 2n\}$ of size n and sends $i, c_B^{i,j} := \text{GenChallenge}_{\mathcal{F}}$ for $j \in \mathcal{J}^i$ to the Seller;
  - For $j \in \mathcal{J}^i$, the Seller opens his commitment to $S^{i,j}$ and checks that $c_B^{i,j} = c_k^{i,j}$ for $k = 1$ or $k = 2$. He opens the commitments to $\Gamma_k^{i,j}$ for only this k;
  - For $j \notin \mathcal{J}^i$, the Seller opens his commitments to $\Gamma_k^{i,j}$ for $\text{k} = 1,2$;
  - The Buyer verifies all the commitments;
  - For $j \in \mathcal{J}^i$, the Buyer decrypts $(c^{i,j}, r^{i,j}) = \text{Dec}_{S^{i,j}}(\gamma_k^{i,j})$, and checks that $c^{i,j} = c_B^{i,j}$ and $\text{VerifyResponse}_{\mathcal{F}}\left(B_{\mathcal{F}}^{i,j}, c_B^{i,j}, r^{i,j}\right) = true$.
- **Stage 4:**
  - The Buyer broadcasts T_1 and T'_1, and then the two parties wait until they are confirmed finally;
  - If transactions T_1 and T'_1 are valid as expected, the Seller broadcasts T_2 and uses signatures $\hat{\sigma}_1, \dots, \hat{\sigma}_a$ to get her payment, and otherwise he broadcasts T'_3 to has his deposit back;
  - If transaction T_2 is valid as expected, the Buyer uses signature $\hat{\sigma}_i$ to calculate secrets $S^{i,j}$ and then decrypts all the values $\gamma^{i,j}$ to get the corresponding challenges $c_k^{i,j}$and responses $r_k^{i,j}$. At the end, the Buyer uses any pair of responses to calculate $x' = \text{Extract}_{\mathcal{F}}(B_{\mathcal{F}}^{i,j}, c_1^{i,j}, r_1^{i,j}, c_2^{i,j}, r_2^{i,j})$. Otherwise, he force-opens time-locked puzzles Φ_i $(i \in \{1, \dots, a \backslash J\})$, and uses any of the opened values d_S^i to get his funds back and obtain the Seller's deposit as well.

Figure 1. The SellWitnessProtocol.

5. CONCLUSIONS

In this paper, with respect to the unfairness between the buyer and seller in the Zero-Knowledge Contingent Payment protocol, we present protocol SellWitness to achieve that, if the seller is malicious, then the honest buyer will get back his payment and some penalty paid by the seller. In the further work, we will optimize SellWitness, e.g., replacing the time-lock building block with some timeless primitives to allow the parties to have unlimited number of valid transactions.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their valuable comments. This work is supported in part by the National Key R&D Program of China (No. 2020YFB1005801) and in part by the National Natural Science Foundation of China (No. 62172396).

REFERENCES

[1] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[J]. Decentralized Business Review, 2008: 21260.

[2] Campanelli M, Gennaro R, Goldfeder S, et al. Zero-knowledge contingent payments revisited: Attacks and payments for services[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017: 229-243.

[3] Greg Maxwell. “The first successful Zero-Knowledge ContingentPayment.”https://bitcoincore.org/en/2016

[4] Banasik W, Dziembowski S, Malinowski D. Efficient zero-knowledge contingent payments in cryptocurrencies without scripts[C]//European symposium on research in computer security. Springer, Cham, 2016: 261-280.

[5] Bentov I, Kumaresan R. How to use bitcoin to design fair protocols[C]//Annual Cryptology Conference. Springer, Berlin, Heidelberg, 2014: 421-439.

[6] Andrychowicz M, Dziembowski S, Malinowski D, et al. Fair two-party computations via bitcoin deposits[C]//International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg, 2014: 105-121.

[7] Goldrcich O, Vainish R. How to solve any protocol problem-an efficiency improvement[C]// Conference on the Theory and Application of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1987: 73-86.

[8] An J H, Dodis Y, Rabin T. On the security of joint signature and encryption[C]//International conference on the theory and applications of cryptographic techniques. Springer, Berlin, Heidelberg, 2002: 83-107.

[9] Boneh D, Shen E, Waters B. Strongly unforgeable signatures based on computational Diffie-Hellman[C]//International Workshop on Public Key Cryptography. Springer, Berlin, Heidelberg, 2006: 229-240.

[10] Katz J, Lindell Y. Introduction to modern cryptography[M]. CRC press, 2020.

[11] Paillier P. Public-key cryptosystems based on composite degree residuosity classes[C]//International conference on the theory and applications of cryptographic techniques. Springer, Berlin, Heidelberg, 1999: 223-238.

[12] Boneh D, Naor M. Timed commitments[C]//Annual international cryptology conference. Springer, Berlin, Heidelberg, 2000: 236-254.

[13] Rivest R L, Shamir A, Wagner D A. Time-lock puzzles and timed-release crypto[J]. 1996.

[14] Bellare M, Goldreich O. On defining proofs of knowledge[C]//Annual International Cryptology Conference. Springer, Berlin, Heidelberg, 1992: 390-420.

[15] Goldreich, O.: Foundations of Cryptography, vol. 1. Cambridge University Press, New York (2006). ISBN: 0521035368.

[16] Goldwasser S, Micali S, Rackoff C. The knowledge complexity of interactive proof-systems[M]//Providing Sound Foundations for Cryptography: On the Work of Shafi Goldwasser and Silvio Micali. 2019: 203-225.

AUTHORS

Peifang Ni received the Ph.D. degree in the Institute of Information Engineering, Chinese Academy of Sciences, in 2020. She is currently a postdoctoral with the Trusted Computing and Information Assurance Laboratory, Institute of Software, Chinese Academy of Sciences. Her major research interests include applied cryptography, security protocol and blockchain consensus.

MoodLink: A Data-Driven Social Interactive Mobile Application for Depression Relief using Artificial Intelligence and Natural Language Processing

Yilan Zhao[1] and Yu Sun[2]

[1]Irvine High School, 4321 Walnut Ave, Irvine, CA 92604
[2]California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

Abstract

As adolescent suicide rates grew significantly in the past decade, depression, anxiety, and other mental disorders were largely held responsible for the growth [9]. However, these medical conditions are often overlooked during their early stages where symptoms are still remediable. Delayed or inattentive response to address the issue usually results in higher suicides rates or in lesser cases, mental ailments carried into adulthood. In an attempt to remedy the mental health crisis, countless mental health interventions are being introduced as means to mitigate the circumstances. In this project, we developed a mobile application that serves as a comprehensive therapy—journal and group therapy—for those struggling with mild to moderate depressive symptoms [10]. The application utilizes both the Sentimental AI and natural language processing in its backend server to generate accurate matches of users who share similar struggles, allowing users to connect and resonate with each other emotionally [11]. The application also provides a private and safe space for users to openly express their thoughts, alleviating their stress through daily journal entries.

Keywords

Machine learning, Flutter, Adolescent Mental Health, Depression.

1. Introduction

A crucial stage of a person's mental wellness is developed during the period of adolescence where healthy social and emotional habits are formed. Currently, however, studies reported that one in seven 10-19 year-olds are struggling with at least one mental disorder. Among teenagers aged 15-19 years old, the fourth leading cause of death is suicide. Suicide rates have increased 30% in the past 20 years due to a variety of causes, such as mental illnesses and disabilities, most notably anxiety and depression [1]. A failure to address these mental health conditions among adolescents results in these ailments continuing into adulthood, as evidenced by Fombonne et al [2]. This causes a significant impairment in physical and mental faculties, becoming a contributor to growing suicide rates among adults as well. To remedy this growing mental health crisis among adolescents, mental health awareness, promotion, and prevention techniques are a key factor. Adolescents should be taught how to best regulate their emotions, build mental and

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 69-79, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121606

emotional resilience, and most importantly create and maintain positive and supportive social environments and networks.

Several techniques have been proposed in the past decades to deal with this mental health crisis, to varying levels of effectiveness. Antidepressants, for instance, saw clinical introduction in the 1950s. The drug sought to increase the activity of neurotransmitters, such as serotonin, dopamine, and norepinephrine, and help lessen depression and anxiety symptoms in patients. The usage has shown to improve a patient's mood and emotions, while also increasing one’s appetite and concentration. However, while the drug improved depressive symptoms, side effects were observed: some had to be discontinued for certain patients due to insomnia, dry mouth, or nausea [3].

Online support groups, on the other hand, emerged as the new form of mental health counseling, allowing individuals to more conveniently express their struggles and feelings through virtual platforms. Even without face-to-face interaction with the other individuals, the method proved to be a source of individual empowerment and collective group identity. However, this effectiveness can occasionally be limited, as Wentzer and Bygholm stated in their study: online support groups showed improvement “not for collective empowerment” [4].

With the rise of new social media platforms, some tried to extend public engagement about mental health issues by heavily utilizing hashtags on platforms, such as Instagram. Mental health awareness campaigns across many social media platforms are conducted annually. Instead of finding how the public engages collectively and building a coherent sense of community concerned with mental illnesses, McCosker and Gerrard found that only 15% of users in these campaigns were considered legitimately depressed [5]. The rest, they found, had profile activity indicative of socially engineered depression made namely to garner attention. Only a minority of social media posts in the data set revealed real signs of mental struggles, usually lacking heavy use of hashtags. This discovery demonstrated that social media have the potential to demean the purpose of those who are in fact experiencing depressive symptoms.

Our approach to this method is a lightweight social media engineered to fit the needs of those with depression and anxiety [15]. This application combines multiple aspects of previously stated approaches, namely journaling as well as mental health awareness through social media interactions. The end goal of this application is to allow users to express frustrations and mental struggles through both private and public interactions with other like-minded users. Users will be able to discuss their own mental state through an electronic journaling system. On the public interaction side, users are matched up through a sentiment analysis algorithm, which will recommend users based on shared content between sets of posts. We believe that matching users up and letting them communicate and understand each other will help promote an inclusive and supportive environment.

In two application scenarios, we demonstrate how the above combination of techniques The rest of the paper is organized as follows: Section 2 discusses the challenges in this research as well as during the experimentation phase of our AI; Section 3 describes our methodology and solution to those challenges through a mobile application and backend server; Section 4 displays the experiments we completed to test the accuracy of the AI backend server we created; Section 5 compares our solution to other app-based mental health interventions and experiments. Finally, Section 6 concludes the study and states future work of this project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. The Complexity of Partner Recommendation Systems and Sentimental Analysis

Traditional social media's friend recommendation is often based on various factors, such as: the user's phone contact, search history, mutual friends, followings [12]. However, for the specific purpose of this application, friends' recommendations need to be tailored to the user's circumstances. As one major goal of this application is to pair up users who share similar emotions or experiences so that they can resonate with each other, the application needs to use a backend AI—the Sentimental AI—to make accurate match-ups based on the tone and content of the user's journal entries. However, since most users of this app are expected to have negative behavior, the effectiveness of using a sentimentality analysis AI is limited, as most posts would be registered as sad. Natural language processing is another option to consider, which analyzes content instead of emotion. Natural language processing will be able to grasp certain subjects such as interests, passions, and frustrations to a given level of specificity.

2.2. Compromises Between User Privacy and Robust User Interaction

In order to create a safe, personal space for users to freely express their emotions without the concerns for others' comments, the application will need a degree of user privacy. However, one potential issue for a system such as this is that users may not receive any attention if their account is too private. Inversely, users may also receive unwanted attention if they make their account public. Reaching a balance or a compromise between these two situations can be tricky to accomplish. For an app centered around privacy, it can be difficult to determine the length to which users can interact with each other. Users cannot simply befriend and start conversations with everyone from their recommended friend list because that becomes an invasion of other's privacy. Likewise, users should not be able to see others' posts should they contain private information. Hence, in order to start a conversation, the user needs to send a friend request that needs to be manually accepted by the other user in order for the conversation to begin. Guaranteeing this level of user privacy, the application helps to encourage a safe space and a positive environment.

2.3. User Experience and Associated Ideal Use Case with User Retention in Mind

In order to maintain a positive user experience, the application will need to include an intuitive user interface. To avoid users finding the application confusing to navigate due to over - complicated app features, the interface will require navigation bars that help segment the application into simple chunks. Intrusive and useless features, such as ads and constant notifications, will need to be eliminated. Instead, features that meet user needs, such as communication and posts, will be preserved to create a simple, clean app design that improves user experience. The intended use case of the app must also be taken into consideration. Users ideally should use this application in short bursts during times of stress or generally semi-daily use. Users should spend half of their time journaling and the other half of their time chatting with others. The app must be designed with this use case in mind to keep user retention high. Otherwise users may grow overwhelmed or impatient within the first couple days of use and give up on using it.

3. SOLUTION

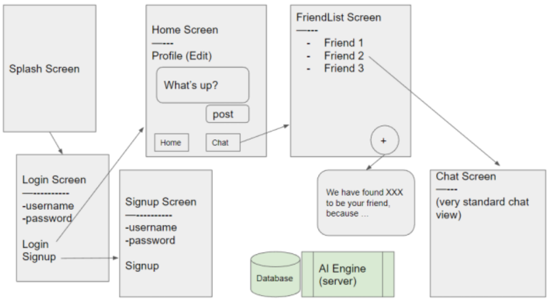


Figure 1. Overview of the solution

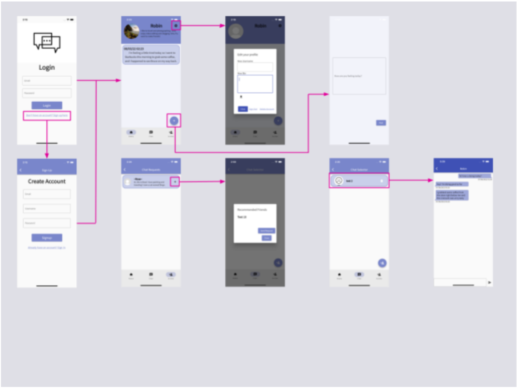


Figure 2. Screenshot of using process

3 Purposes

- allow user alleviate stress through journal entries by offering them a private place to express their emotions
- accurately match users who share similar emotions/experiences by analyzing journal entries using sentimental AI
- allow users to communicate and resonant with each other (mimics self-help group in therapy)

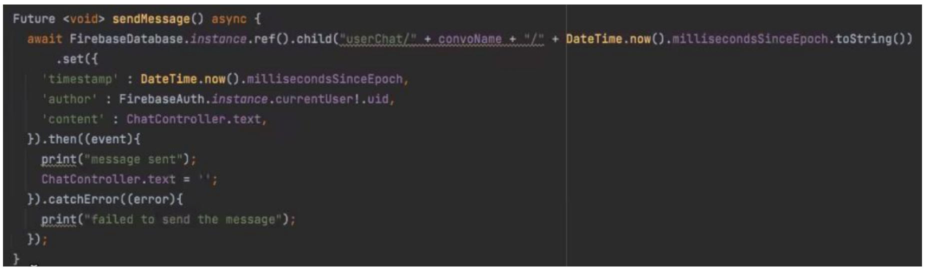
3 Components

- frontend application (Dart in Flutter)
    - 3 sections:
        - Journal entry page
        - recommended friends page
        - chat page
- backend cloud database (Firebase)
    - Realtime database
    - storage
    - authentication
- backend machine learning server (Sentimental AI written in Python)

- access data gathered to analyze the emotional tone and match users who have similar emotions

This application provides a safe, personal space to help users alleviate stress through journal entries and accurately matches up users who share similar emotions or experiences so that they can resonate with each other. The application relies upon three central components: the frontend application, which is done through Flutter; the backend cloud database, which is done through Firebase, and the backend machine learning server, which assists in user recommendation.

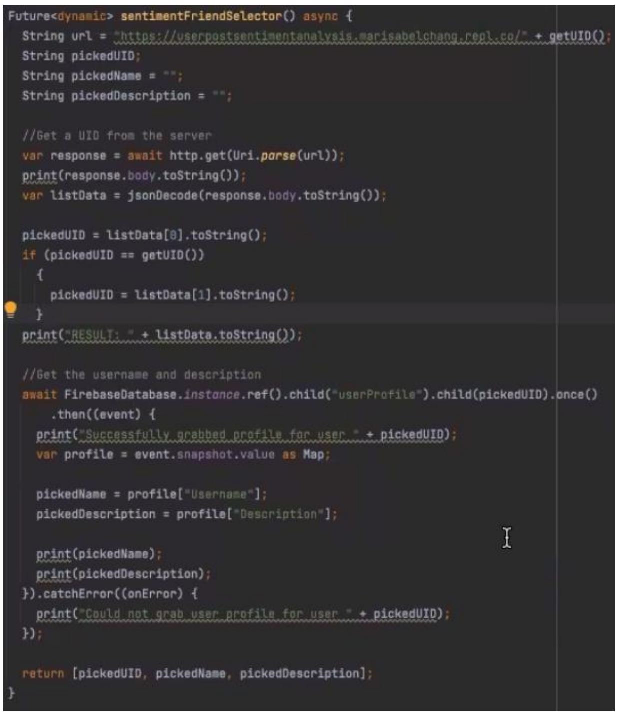
The frontend is only accessible if the user is authenticated by Firebase's Authentication service [13]. This is done through a simple login system. To sign up, users will only have to provide an email address and a password. The frontend itself consists of a journal entry section, a friend recommendation section, and a chat section. On the user's profile page, they can view their entries from most recent, as well as change their profile info, such as user name, description, and profile picture. Users will have to write at least one post in their journal in order to access the recommendation system in the friends page. This page relies on a backend machine learning server which will use sentiment analysis to determine which user is most similar to the current user based on post content. The resulting recommendation is shown to the user for them to send an invite if they would like. After an invite is sent and the other user accepts the friend request from their end, then both users are able to communicate in a chat screen which functions similarly to standard chat screens present in applications such as WhatsApp and WeChat. User profile information, user journals, and chat logs are all stored within Firebase's Realtime Database service with appropriate rules to ensure privacy.



```
Future <void> sendMessage() async {
  await FirebaseDatabase.instance.ref().child("userChat/" + convoName + "/" + DateTime.now().millisecondsSinceEpoch.toString())
      .set({
    'timestamp' : DateTime.now().millisecondsSinceEpoch,
    'author' : FirebaseAuth.instance.currentUser!.uid,
    'content' : ChatController.text,
  }).then((event){
    print("message sent");
    ChatController.text = '';
  }).catchError((error){
    print("failed to send the message");
  });
}
```

Figure 3. Screenshot of code 1

For this specific project, Google Firebase is used as a platform that offers an active backend for analytic, authentication, databases, fire storage and more. Its Realtime Database service enables the application to store and sync data between users in real time, consisting of five subdirectories: userChat, userFriend, userInvite, userPost, and userProfile. Each subdirectory has its own properties, some of which have their own sub-collections. For example, userChat stores chat content by organizing each conversation under a key name that is a combination of each users' User Identification code (UIDS). Then under each ConvoName, a subcategory is created using the timestamp the message is sent. This organization based on timestamp allows the application to display messages by its chronological order on the chat screen. The timestamp category then contains the actual content of the message, its author, and its timestamp so that when data is queried, the application displays messages based on the chronological sender and time. A successful message will be printed in the console if the message is recorded in the realtime database; if not, a failure message will be shown. By storing the conversation messages into the realtime database in an organized manner, the application can extract the information more efficiently and display them logically. The above piece of code is meant for sending messages. It creates a new list of key/values pairs within the Realtime Database containing all the appropriate metadata per message. The address of this message is within userChats, within the corresponding convoName.



```
Future<dynamic> sentimentFriendSelector() async {
  String url = "https://userpostsentimentanalysis.marisabelchang.repl.co/" + getUID();
  String pickedUID;
  String pickedName = "";
  String pickedDescription = "";

  //Get a UID from the server
  var response = await http.get(Uri.parse(url));
  print(response.body.toString());
  var listData = jsonDecode(response.body.toString());

  pickedUID = listData[0].toString();
  if (pickedUID == getUID())
    {
      pickedUID = listData[1].toString();
    }
  print("RESULT: " + listData.toString());

  //Get the username and description
  await FirebaseDatabase.instance.ref().child("userProfile").child(pickedUID).once()
      .then((event) {
    print("Successfully grabbed profile for user " + pickedUID);
    var profile = event.snapshot.value as Map;

    pickedName = profile["Username"];
    pickedDescription = profile["Description"];

    print(pickedName);
    print(pickedDescription);
  }).catchError((onError) {
    print("Could not grab user profile for user " + pickedUID);
  });

  return [pickedUID, pickedName, pickedDescription];
}
```

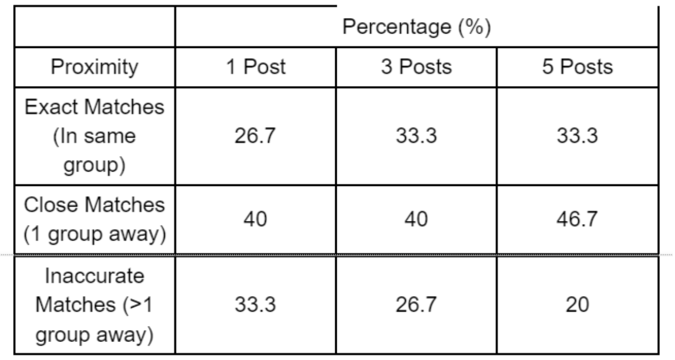
Figure 4. Screenshot of code 2

This application's backend consists of both the Google Firebase and the Sentimental AI hosted by a Python server that is running flask. To structure an accurate friend recommendation system, the Firebase and the Python server consistently communicate with each other. Because the user data is stored in the Firebase, the Python server will have to access and read data from the database before producing the analyzed results. Sentimental Analysis is a type of machine learning algorithm that can calculate the "polarity" of a given piece of text. A common use is to determine user satisfaction. A specification of sentimental analysis, called Natural Language Processing, can be used to determine the topic discussed instead of emotion, though it is still calculated as a polarity score. Using Flask, a Python server library, the server hosts the Sentimental Analysis AI, which then reads through all posts made by the user and determines its polarity for each post. The AI proceeds to calculate the user's own polarity and selects one user from the database whose polarity best matches the target user. During this process, the AI filters out the user's friends to avoid recommending a friend who already exists in the target user's friend list. After producing the best match, the Python server makes a Firebase call to the Realtime database's userProfile subdirectory in order to display the recommended friend's information to the user.

4. EXPERIMENT

4.1. Experiment 1

In order to evaluate the accuracy of the friend recommendation system, we tested the Sentimental AI in various scenarios. For this experiment, we utilized synthetic data, which were 15 test subjects divided into five groups ranging from experiencing negative, neutral, to positive emotions. The scale for these emotions were chosen arbitrarily. Subjects 1, 2, and 3 from Group 1 exhibit extreme depressive symptoms while Subjects 4, 5, and 6 from Group 2 exhibit moderate depressive symptoms. The inverse characteristic occurs for groups 4 and 5. Group 3 is considered to be neutral. One journal entry displaying the corresponding emotion will be posted by each test subject and stored into the Google Firebase. The Sentimental AI will then proceed to make friend recommendations based on the tone and content analyzed from each journal entry, and matching results for each test subject will be recorded. Data is processed according to the proximity of the recommendation — recommendations of members within a group are considered "exact." Recommendations for a member one group away (e.g. a Group 2 member being recommended a Group 1 member) is considered "close." Anything else will be considered an inaccurate recommendation.



	Percentage (%)		
Proximity	1 Post	3 Posts	5 Posts
Exact Matches (In same group)	26.7	33.3	33.3
Close Matches (1 group away)	40	40	46.7
Inaccurate Matches (>1 group away)	33.3	26.7	20

Figure 5. Table of proximity vs accuracy

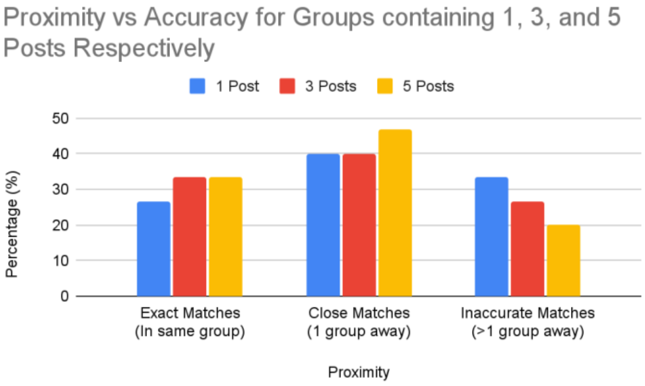
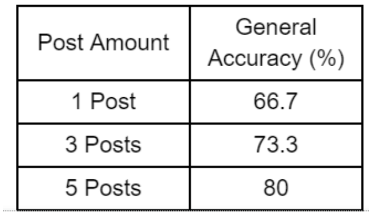


Figure 6. Graph of proximity vs accuracy



Post Amount	General Accuracy (%)
1 Post	66.7
3 Posts	73.3
5 Posts	80

Figure 7. Table of post amount vs general accuracy

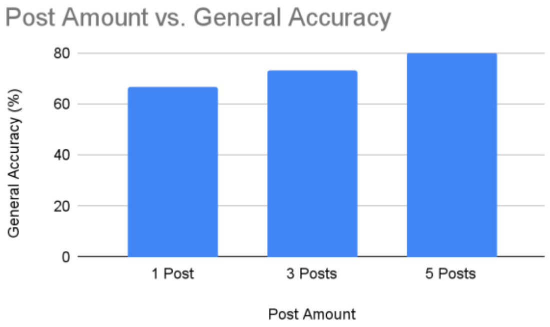
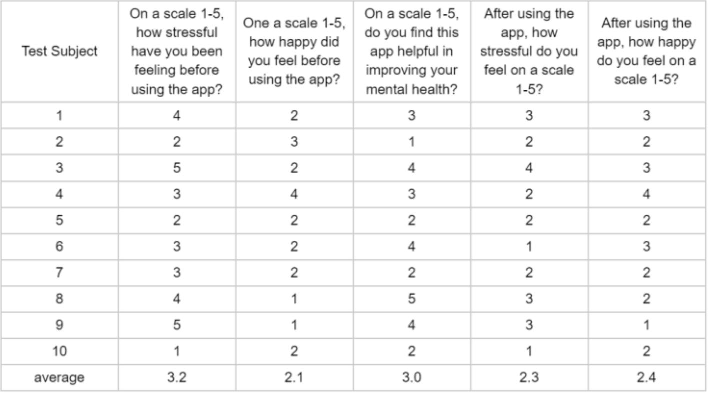


Figure 8. Graph of post amount vs general accuracy

The results indicate several interesting conclusions. Trials for 3 and 5 posts had the highest frequency of exact matches, at 33.3% each. The 1 post trial had the worst accuracy of exact matches, at 26.7%. Trials for 1 post and 3 posts tied in frequency with regards to close matches at 40%. The 5 post trials scored the highest frequency for close matches at 46.7%. For inaccurate matches, the 1 post trial scored the worst, at 33.3% inaccuracy. This is followed up by the 3 post trial, at 26.7% inaccuracy, and the 5 post trial, at 20% inaccuracy. As such, accuracy in this experiment can also be said to be generally proportional to the amount of posts that users made. When determining overall accuracy by combining the frequency of exact and close matches as a percentage, the 1 post trial was the least accurate at 66.7% accurate, the 3 post trial at 73.3%, and the 5 post trial at 80% accurate. Regardless of trial, however, close matches were the most frequent recommendation.

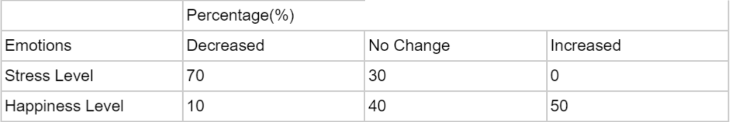
4.2. Experiment 2

In addition to checking the artificial intelligence's accuracy, it is also important for us to check if our implementation of therapy methodologies (including our artificial intelligence) is successful or not. To do this, we designed another experiment that will look at user happiness and stress level prior to using the app and the resultant levels after using the application. 10 test subjects for this experiment were given a simple questionnaire after testing out the application. The questionnaire asked participants questions pertaining to stress level and satisfaction with the app. Participants were asked to answer such questions from a grading scale of 1-5.



Test Subject	On a scale 1-5, how stressful have you been feeling before using the app?	One a scale 1-5, how happy did you feel before using the app?	On a scale 1-5, do you find this app helpful in improving your mental health?	After using the app, how stressful do you feel on a scale 1-5?	After using the app, how happy do you feel on a scale 1-5?
1	4	2	3	3	3
2	2	3	1	2	2
3	5	2	4	4	3
4	3	4	3	2	4
5	2	2	2	2	2
6	3	2	4	1	3
7	3	2	2	2	2
8	4	1	5	3	2
9	5	1	4	3	1
10	1	2	2	1	2
average	3.2	2.1	3.0	2.3	2.4

Figure 9. Table of test subject



	Percentage(%)		
Emotions	Decreased	No Change	Increased
Stress Level	70	30	0
Happiness Level	10	40	50

Figure 10. Table of emotions change

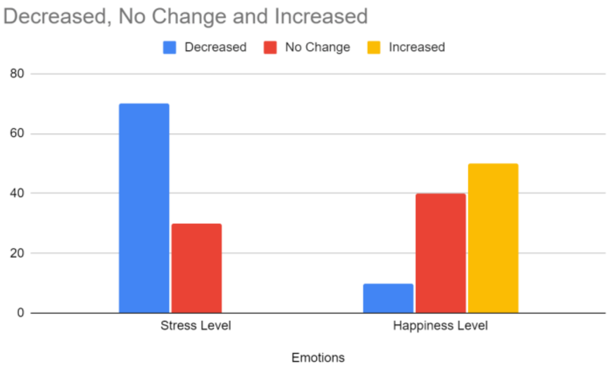


Figure 11. Graph of Emotion change

The table and the graph display a few interesting trends. The test subjects responded most positively to the decrease in stress level, as reflected by a reduced average stress rating from 3.2 to 2.3 after the usage of the application. The graph also demonstrates this trait as no test subjects are shown to report an increase in stress. On the other hand, happiness level sees an increase from an average of 2.1 to 2.4; however, 10% of the test subjects stated a decrease in happiness level while 40% reported no change. This may be caused by the test subject’ s unfamiliarity with the app features as they were still adjusting to the new form of mental health intervention. Overall, the test subjects rated 3.0 on the effectiveness of the application on their mental health: 70% reported reduced stress, and 50% displayed improved happiness.

Both experiments yielded moderately positive results, exhibiting a general success of the application. Experiment 4.1 saw that the accuracy of the friend recommendation system is directly proportional to the amount of journal entries an user has: the highest frequency of exact matches was with the 5 post trials at 46.7%. The experiment also concluded that if a user has five entries, the AI can generally guarantee 80% matches to fall within at most one group of the target users. Experiment 4.2 displayed 70% reduced stress level and 50% increased happiness level among test subjects after the usage of the application. However, 10% of subjects reported decreased happiness levels. The subjects’ unfamiliarity with the app features might have contributed to the negative feedback.

5. Related Work

Mohr et al. for this study performed an extensive coaching experiment using their own app suite called Intellicare [6]. Unlike our application, Intellicare has multiple applications instead of just one. Researchers found that bundling several apps into a comprehensive self-care package was successful in alleviating depression and anxiety (Mohr et al, 2017). Intellicare is also focused more on active self-betterment and coaching rather than our application, which focuses more on user interactions and journaling. According to the researchers, “...apps tend to use simple interactions, are quick to use, and support a single or limited set of related tasks" (Mohr et al, 2017). Our application aims for more engagement than theirs by acting as a chat discussion app as well as an app to privately post and use as a safe space. Researchers for this study also extensively used coaching as an experiment method (Mohr et al, 2017). Our application similarly emphasizes constant communication and assistance, as it is built into the app's infrastructure.

Birney et al. propose a mobile web app called MoodHacker, which utilizes cognitive behavioral therapy skills, to help reduce negative cognitions for working adults who suffer from depression [7]. The app produced significant results: improvement on testing subjects' depression symptoms and other workplace-induced negative cognitions (Birney et al, 2016). Instead of focusing on adolescent's mental health related issues, this paper aims to develop a mobile application catered towards working adults in the United States, who amount for an estimated $210.5 billion in productivity loss (Birney et al, 2016). According to the researchers, the mobile web app seeks to "develop effective interventions that can be more widely disseminated … and directly to individuals who will not seek face-to-face care" (Birney et al, 2016). Our application shares a similar objective: allowing adolescents who suffer from mental health issues to communicate and resonate with each other in a virtual setting where in-person interaction is not needed.

Fuller-Tyszkiewicz et al. seeks to evaluate the usability of a mental health application by having three groups—diagnosed individuals who suffer from depression, professionals such as doctors, and specialized mental health researchers—-to rate the application in terms of its quantitative and qualitative impact on its users [8]. The study found that the two expert groups favored app features that emphasize self-betterment and serve as cognitive treatment for the users but showed their concern for the challenging self-navigation of the application (Tyszkiewicz et al, 2018). This paper highlights the importance of the functionality of mental health applications by implementing simple self-navigation features (Tyszkiewicz et al, 2018). Our application reflects this by setting up an intuitive user interface that contains only necessary features organized in a coherent fashion, greatly optimizing user experience by eliminating overcomplicated app design. According to the researchers, the usability of the application is vital because "perceptions of navigability and quality of content are likely to impact participant engagement and treatment compliance" (Tyszkiewicz et al, 2018).

6. CONCLUSIONS

In summary, our method was to develop a mobile application which serves as a comprehensive therapy application for those suffering from depression [14]. This application is connected to a backend server that utilizes natural language processing to recommend people to those that share their interests and struggles. This Sentimental AI makes friend recommendations by accessing data from the cloud database, the Google Firebase, and analyzing the content and emotional tone. The app also includes journal therapy and group therapy techniques within its features. The application seeks to alleviate stress by providing a safe, private space for users to freely express their thoughts and emotions while also allowing users to resonate with one another. Two experiments were conducted to determine the objective and subjective accuracy of the backend AI server. The first experiment utilized fabricated journal entries created by users— who showed varying levels of emotions—and recorded the matching results generated by the Sentimental AI. The experiment results supported the fact that matching accuracy is usually directly proportional to the number of entries an user made: 5 posts revealed 80% general accuracy while 1 post was only 66% accurate. The questionnaire revealed that 80% of users felt less stressed and 50% of users felt happier after using the app. These results indicate that the AI is at least both generally accurate and effective at bolstering user interaction and satisfaction. The experiments conducted solved challenges brought up prior by both analyzing potential concerns with the accuracy of the artificial intelligence backend but also with the user experience of the app in relation to the AI system implemented.

As evidenced by our experiments on AI accuracy, the accuracy of the natural language processing algorithm used in our backend server could be better optimized and more robust. It is able to pick up on general topics but can struggle when given posts with minimal data and also struggles with picking up on fine details. The AI recommendation system as a whole may also not be utilized well enough by the user base if they do not consider the application to be useful enough. In regards to the app's frontend, more quality of life features could also be added to the application to better ensure user satisfaction when it comes to user experience.

As we collect more real data from the influx of new users, further experimentation will be conducted to improve the accuracy of the Sentimental AI. Its limitation will be addressed as we explore the possibility of increasing the AI's capability of analyzing content in more details when it is given minimal data. Moreover, updates on the application itself will be added in an attempt to improve user experience. We hope to provide users not only an effective but also friendly app.

REFERENCES

[1] Kieling, Christian, et al. "Child and adolescent mental health worldwide: evidence for action." The Lancet 378.9801 (2011): 1515-1525.

[2] Fombonne, Eric, et al. "The Maudsley long-term follow-up of child and adolescent depression: I. Psychiatric outcomes in adulthood." The British Journal of Psychiatry 179.3 (2001): 210-217.

[3] Paolucci, Stefano, et al. "Post-stroke depression, antidepressant treatment and rehabilitation results." Cerebrovascular diseases 12.3 (2001): 264-271.

[4] Wentzer, Helle S., and Ann Bygholm. "Narratives of empowerment and compliance: studies of communication in online patient support groups." International journal of medical informatics 82.12 (2013): e386-e394.

[5] McCosker, Anthony, and Ysabel Gerrard. "Hashtagging depression on Instagram: Towards a more inclusive mental health research methodology." new media & society 23.7 (2021): 1899-1919.

[6] Mohr, David C., et al. "IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety." Journal of medical Internet research 19.1 (2017): e6645.

[7] Birney, Amelia J., et al. "MoodHacker mobile web app with email for adults to self-manage mild-to-moderate depression: randomized controlled trial." JMIR mHealth and uHealth 4.1 (2016): e4231.

[8] Fuller-Tyszkiewicz, Matthew, et al. "A mobile app–based intervention for depression: end-user and expert usability testing study." JMIR mental health 5.3 (2018): e9445.

[9] Novick, Lloyd F., Donald A. Cibula, and Sally M. Sutphen. "Adolescent suicide prevention." American journal of preventive medicine 24.4 (2003): 150-156.

[10] McCurdy, Ashley P., et al. "Effects of exercise on mild-to-moderate depressive symptoms in the postpartum period: a meta-analysis." Obstetrics & Gynecology 129.6 (2017): 1087-1097.

[11] Bangare, S. L., et al. "Using Node. Js to build high speed and scalable backend database server." Proc. NCPCI. Conf. Vol. 2016. 2016.

[12] Kaplan, Andreas M., and Michael Haenlein. "Users of the world, unite! The challenges and opportunities of Social Media." Business horizons 53.1 (2010): 59-68.

[13] Khawas, Chunnu, and Pritam Shah. "Application of firebase in android app development-a study." International Journal of Computer Applications 179.46 (2018): 49-53.

[14] Teng, Chia-Chi, and Richard Helps. "Mobile application development: Essential new directions for IT." 2010 Seventh International Conference on Information Technology: New Generations. IEEE, 2010.

[15] Dobson, Keith S. "The relationship between anxiety and depression." Clinical Psychology Review 5.4 (1985): 307-324.

USER REPAIRABLE AND CUSTOMIZABLE BUZZER SYSTEM USING MACHINE LEARNING AND IOT SYSTEM

Leheng Huang[1] and Yu Sun[2]

[1]Arcadia High School, 180 Campus Dr, Arcadia, CA 91006
[2]California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

ABSTRACT

The creation and sustainability of academic teams have long been unnecessarily difficult due to the exorbitant costs of purchasing and maintaining equipment [1][2]. These costs serve as a major barrier, especially in poorer areas where securing the funds for this equipment is difficult [3]. In addition, when the equipment eventually breaks, it is often difficult to repair, forcing academic teams to purchase a new set of equipment. This project attempts to provide a product that can drastically lower the equipment's costs and allow the user to modify and repair it as necessary. This project resulted in the development of the Argo Buzzer System which was created with input from experienced academic team members and it has proven that it is comparable to modern buzzer systems for a fraction of the cost [4].

KEYWORDS

Electronics, Machine learning, IoT system.

1. INTRODUCTION

Academic teams are an important aspect of school environments by providing a designated area for like-minded scholars to congregate and challenge themselves with real-world problems [5]. Locally, this will often create an environment that stimulates the exchange of ideas and collective growth of knowledge of the individual team. On a regional scale, academic teams and their associated competitions allow scholars to gauge their level of understanding relative to others in addition to promoting the meeting of individuals who are interested in similar topics [6]. Overall academic teams are vital to the positive development of scholarly students towards their academic goals. However, around the world, many academic teams struggle with financial limitations due to the excessively high cost of equipment and maintenance [7]. The issue is disproportionately disadvantaged towards poorer communities where many passionate students are unable to secure funding from their schools to establish these teams.

Currently, the main source of buzzers is made by a few companies [8]. However these systems often cost an exorbitant amount and once broken, are near impossible for the user to repair. The hefty procurement cost is often a big initial obstacle for those looking to start an academic team, and the recurring financial obstacle of purchasing a new system when the old one breaks is a constant threat to the existence of academic teams.

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 81-88, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121607

The method employed in this research paper was the design and construction of a new buzzer set. The design and construction of the new system adheres to the priorities on the academic teams and as a result features unparalleled user-friendliness and repairability [9]. This feature is seldom found on any other buzzer set if it can be found at all. In addition, the new system utilizes cost effective parts designed to lower cost while maintaining overall integrity of the system.

As part of multiple of the aforementioned academic teams, I was able to gather the opinions of many experienced and esteemed members of the community. Guided by their constructive criticism and factual evidence gathered through various means, this paper, and the associated patent seeks to provide an alternative by designing a product using cheap parts that can be easily sourced, assembled, and modified [6]. In addition to getting feedback from academic teams, I also conducted multiple tests on the buzzer system to assure the proper functioning of the system. During the tests, the buzzer pressed and the reset method was alterensnated multiple times to ensure that all systems functioned as designed.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we met during the experiment and designing the sample; Section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 2; Section 4 presents the relevant details about the experiment we did, following by presenting the related work in Section 5. Finally, Section 6 gives the conclusion remarks, as well as pointing out the future work of this project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Configuring the Product

A significant challenge that I faced during the creation of the product was determining the exact configuration of the machine. Each academic team had its own requirements and preferences, and sometimes it was hard to differentiate which was a requirement and which were not required but strongly preferred. In addition to these considerations had to be made for cost, engineering practicality, and personal sanity. Certain preferences would have added too much to the cost thus defeating its purpose, others would have made building and repairing unnecessarily difficult.

2.2. Designing the Exterior

The design of the exterior shell falls into a separate field of study and requires a separate set of skills than designing electronics. Thus it is natural that most people who build electronics struggle to design an aesthetically pleasing and functional shell. The dilemma that confronts most people during design is the choice of the material. The list of possible materials that can be used is usually constricted by a multitude of factors from the environment to user-friendliness etc. From this resulting list, finding a material that balances all the factors is usually very difficult.

2.3. Determining the Activator

Choosing the proper activator is essential when designing buzzer systems because it determines the way the system will be used and its versatility. Most contemporaries feature 3 main types of activators: handheld buzzers, pedal buzzers, or tabletop buzzers. Each has its respective pros and cons, for example, the pedal buzzers allow for foot buzzing, etc.

3. Solution

The Argo Buzzer System is a buzzer system that features unparalleled user-repairability and upgradability. At the heart of the Argo, Buzzer System is the Arduino Mega; the Arduino Mega microcontroller can be easily reprogrammed by the user which allows for very good adaptability [12]. The default system provides 8 simple handheld buzzers which are individually replaceable, in addition, individual buzzers can be removed or added based on the individual needs. The code provides support for up to 20 players, if the capacity is reached the code provides an input pin and an output pin so that 2 systems can be combined for a total capacity of 40 players. When the buzzer is pressed it completes a circuit that the Arduino detects, when the signal is detected it changes a boolean value and an output is sent to the buzzer and corresponding LED. In addition, all other players are locked out disabling their ability to buzz while the player who buzzed answers the question. To reset the system the operator can either manually press the reset button or activate the laser detection system. The button sends a simple signal to the Arduino after which the boolean is reset ending the lockout and resetting the system. The same procedure occurs with the laser. The major components of the system are as follows:

- Arduino Mega Microcontroller and the buzzer and lockout algorithm [13]
- The individual buzzers and their corresponding circuitry
- The reset button and the laser detection system
- The buzzer

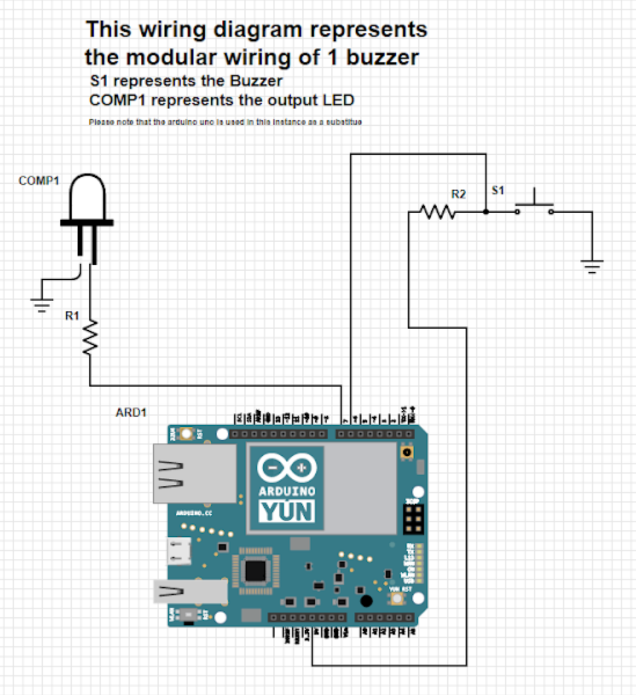
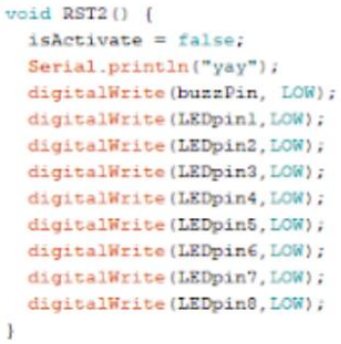
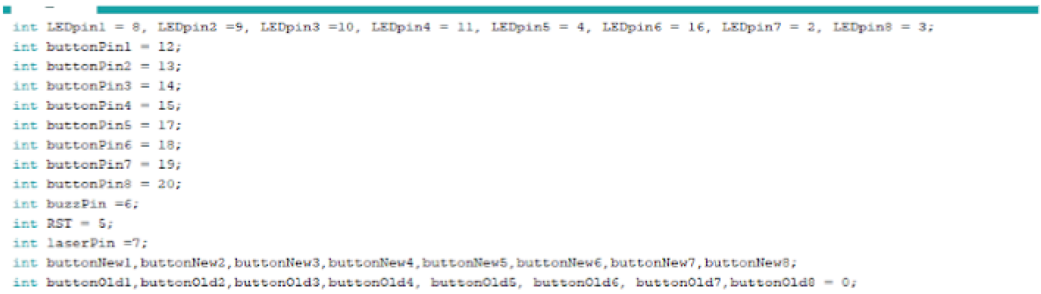

This wiring diagram represents
the modular wiring of 1 buzzer
S1 represents the Buzzer
COMP1 represents the output LED
Please note that the arduino uno is used in this instance as a substitue
COMP1
R1
R2
S1
ARD1
ARDUINO
YUN

Figure 1. Overview of the solution



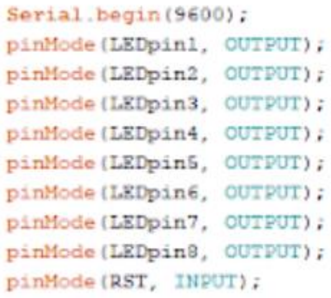
```
void RST2() {
  isActivate = false;
  Serial.println("yay");
  digitalWrite(buzzPin, LOW);
  digitalWrite(LEDpin1,LOW);
  digitalWrite(LEDpin2,LOW);
  digitalWrite(LEDpin3,LOW);
  digitalWrite(LEDpin4,LOW);
  digitalWrite(LEDpin5,LOW);
  digitalWrite(LEDpin6,LOW);
  digitalWrite(LEDpin7,LOW);
  digitalWrite(LEDpin8,LOW);
}
```

Figure 2. Screenshot of code 1



```
int LEDpin1 = 8, LEDpin2 =9, LEDpin3 =10, LEDpin4 = 11, LEDpin5 = 4, LEDpin6 = 16, LEDpin7 = 2, LEDpin8 = 3;
int buttonPin1 = 12;
int buttonPin2 = 13;
int buttonPin3 = 14;
int buttonPin4 = 15;
int buttonPin5 = 17;
int buttonPin6 = 18;
int buttonPin7 = 19;
int buttonPin8 = 20;
int buzzPin =6;
int RST = 5;
int laserPin =7;
int buttonNew1,buttonNew2,buttonNew3,buttonNew4,buttonNew5,buttonNew6,buttonNew7,buttonNew8;
int buttonOld1,buttonOld2,buttonOld3,buttonOld4, buttonOld5, buttonOld6, buttonOld7,buttonOld8 = 0;
```

Figure 3. Screenshot of code 2



```
Serial.begin(9600);
pinMode(LEDpin1, OUTPUT);
pinMode(LEDpin2, OUTPUT);
pinMode(LEDpin3, OUTPUT);
pinMode(LEDpin4, OUTPUT);
pinMode(LEDpin5, OUTPUT);
pinMode(LEDpin6, OUTPUT);
pinMode(LEDpin7, OUTPUT);
pinMode(LEDpin8, OUTPUT);
pinMode(RST, INPUT);
```

Figure 4. Screenshot of code 3

As the centerpiece of the whole system, all inputs and outputs are directly connected to it. To detect the motion of someone pressing the buzzers, the individual buzzers and their corresponding circuits are constantly powered by a 5v current from the Arduino. When the buzzer circuit is open, the signal is read as “1”. When the buzzer circuit is closed, it completes the circuit and the signal reading changes to 0 and then back to 1 as the button is released. When the signal change is detected, boolean changes disable all other buzzers in the system and send a signal to the output pin, while simultaneously triggering the buzzer and activating the LED which corresponds to the buzzer. The reset button operates in a similar manner to the buzzers, it drops the voltage of the system triggering the Arduino. The laser reset system instead simply sends a voltage to the Arduino which the Arduino detects resetting the system.

4. EXPERIMENT

4.1. Experiment 1

The activator is one of the defining traits of any buzzer system as it determines the ways it can be used. The 3 types of buzzers which are widely used by contemporary systems are peddled tabletop, and handheld. However, for this specific instance, peddle-style buzzers are discounted on account of cost, and ergonomics. To determine between the remaining two choices, a survey will be conducted on the following parameters: sample size of 30 taken from players, participants will select either one of the choices or no preference.

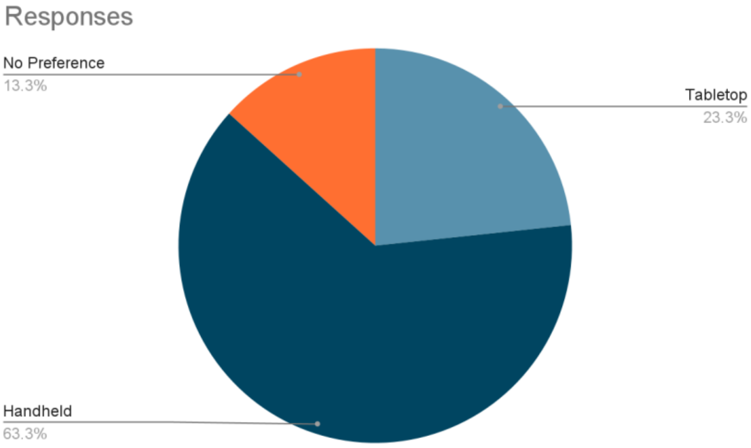
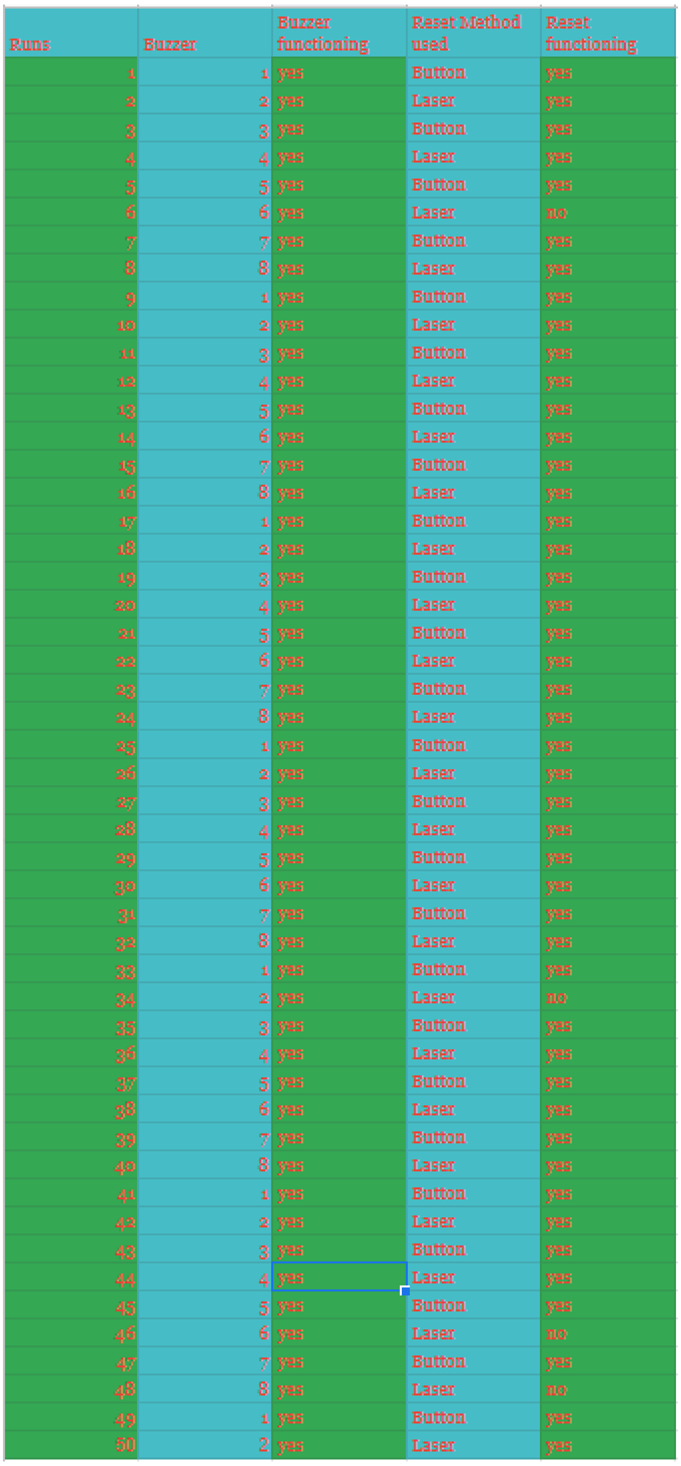


Figure 5. Responses of experiment 1

The result of the experiment showed an overwhelming majority of players preferred hand-held buzzers over tabletop buzzers. This opinion concurs with other surveys conducted online and by other organizations.

4.2. Experiment 2

The functionality of the buzzer system should be comparable to its peers on the market to ensure that academic teams would have all they need. An experiment will be conducted on the following premises to ensure proper functionality: the system will be run 50 times, each handheld buzzer will be cycled through, and the reset method will alternate between the button and the laser activation.



Runs	Buzzer	Buzzer functioning	Reset Method used	Reset functioning
1	1	yes	Button	yes
2	2	yes	Laser	yes
3	3	yes	Button	yes
4	4	yes	Laser	yes
5	5	yes	Button	yes
6	6	yes	Laser	no
7	7	yes	Button	yes
8	8	yes	Laser	yes
9	1	yes	Button	yes
10	2	yes	Laser	yes
11	3	yes	Button	yes
12	4	yes	Laser	yes
13	5	yes	Button	yes
14	6	yes	Laser	yes
15	7	yes	Button	yes
16	8	yes	Laser	yes
17	1	yes	Button	yes
18	2	yes	Laser	yes
19	3	yes	Button	yes
20	4	yes	Laser	yes
21	5	yes	Button	yes
22	6	yes	Laser	yes
23	7	yes	Button	yes
24	8	yes	Laser	yes
25	1	yes	Button	yes
26	2	yes	Laser	yes
27	3	yes	Button	yes
28	4	yes	Laser	yes
29	5	yes	Button	yes
30	6	yes	Laser	yes
31	7	yes	Button	yes
32	8	yes	Laser	yes
33	1	yes	Button	yes
34	2	yes	Laser	no
35	3	yes	Button	yes
36	4	yes	Laser	yes
37	5	yes	Button	yes
38	6	yes	Laser	yes
39	7	yes	Button	yes
40	8	yes	Laser	yes
41	1	yes	Button	yes
42	2	yes	Laser	yes
43	3	yes	Button	yes
44	4	yes	Laser	yes
45	5	yes	Button	yes
46	6	yes	Laser	no
47	7	yes	Button	yes
48	8	yes	Laser	no
49	1	yes	Button	yes
50	2	yes	Laser	yes

Figure 6. Result of experiment 2

During all 50 test runs, the activators all functioned as expected: when pressed, the buzzer buzzed and the corresponding LED lit up. When resetting, the button functioned also as expected, resetting the system when pressed. However, during testing, it was found that the laser reset was occasionally unable to detect when it was being activated due to strong ambient light present during testing. As a result, the laser detection was unable to properly reset the system upon attempted activation.

Experiment 1 regarding the style of the buzzers was within expectations. As a player myself, I preferred handheld buzzers over tabletop buzzers, and many of my fellow players have often complained or expressed dissatisfaction regarding the tabletop buzzers. The experiment confirmed that this was indeed the majority opinion and quantified my suspicions. Experiment 2 regarding the functionality of the buzzer system was within expectation with the exception of the laser reset module. The laser system malfunctioning was a surprise because my previous experience with this particular model of laser module did not have this problem, granted the previous application was used in a dark environment.

5. RELATED WORK

Buzzer systems developed and is currently selling multiple models of their buzzer system [11]. Their models range from 4 players to 16 players with the ability to link up their systems. They sell 2 main types of buzzers, a lockout system, and a special lockout system that shows the orders the players buzz in on. However, their systems are not user customizable or easily repairable as opposed to the Argo Buzzer System. In addition, their systems are ludicrously expensive which are damaging to academic teams.

A similar company by the name of Anderson Enterprises also sells buzzer systems similar to the aforementioned company Buzzer Systems [14]. However, their products suffer from the same issues; poor user repairability and high costs.

Another similar company by the name of zeecraft also sells buzzer systems, their buzzer systems are also similar to the previous companies [15]. However, their buzzer systems are priced at an even more exorbitant amount. In addition, the buzzer system suffers from issues mentioned with the previous buzzer systems.

6. CONCLUSIONS

This paper and its associated product/application were created in response to the growing costs of sustaining and creating an academic team, and to allow for academic teams to upgrade and modify their own buzzer set according to their needs. Through the use of a programmable microcontroller and other easily obtainable parts, the product is completely modular and allows users to repair or modify with basic electronic knowledge and equipment. The design of the product was also designed specifically to suit academic teams and implement feedback given by experienced academic team members as seen in experiment 1. In addition, the product/application has proven that it does function properly with only minor issues that could be easily remedied during experiment 2.

Due to lack of resources, the inputs for requirements, suggestions, and other such opinions were local in scale such the majority opinion may be different in other local areas. The small sample size could also be a liability as it could have inflated or underrepresented certain opinions.

In future, these limitations could be mitigated in the future with larger-scale surveys. In addition, the survey could be improved by asking for more feedback from the participants, especially regarding their reasoning for why they preferred one choice over the other.

REFERENCES

[1] Wilkinson, Adrian, Malcolm Hill, and Paul Gollan. "The sustainability debate." International Journal of Operations & Production Management (2001).

[2] Wadhera, Sidhant. "Exorbitant Costs and Minimal Benefits: the Impact of Hosting the Olympics." Chicago Policy Review (Online) (2020).

[3] Malafeyev, O. A., et al. "The optimization problem of preventive equipment repair planning." AIP Conference Proceedings. Vol. 1978. No. 1. AIP Publishing LLC, 2018.

[4] Tait, Jasmine. "A Comparison of Acoustic Effects of Two Stopper and Crown Systems in the Modern Flute." Canadian Acoustics 29: 40-44.

[5] Gordon, Rick. "Balancing real-world problems with real-world results." Phi Delta Kappan 79.5 (1998): 390.

[6] Dunnell, Robert C., and William S. Dancey. "The siteless survey: a regional scale data collection strategy." Advances in archaeological method and theory. Academic Press, 1983. 267-287.

[7] Jardine, A. K. S., and J. A. Buzacott. "Equipment reliability and maintenance." European Journal of Operational Research 19.3 (1985): 285-296.

[8] Tyson, John J., Katherine C. Chen, and Bela Novak. "Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell." Current opinion in cell biology 15.2 (2003): 221-231.

[9] Darbyshire, Philip. "User-friendliness of computerized information systems." Computers in nursing 18.2 (2000): 93-99.

[10] Abbott, Ann A., and Sharon C. Lyter. "The use of constructive criticism in field supervision." The Clinical Supervisor 17.2 (1999): 43-57.

[11] Robertson, Leon S. "Safety belt use in automobiles with starter-interlock and buzzer-light reminder systems." American Journal of Public Health 65.12 (1975): 1319-1325.

[12] Bolanakis, Dimosthenis E. "A survey of research in microcontroller education." IEEE Revista Iberoamericana de Tecnologias del Aprendizaje 14.2 (2019): 50-57.

[13] Tazi, Imam, Kuwat Triyana, and Dwi Siswanta. "A novel Arduino Mega 2560 microcontroller-based electronic tongue for dairy product classification." AIP Conference Proceedings. Vol. 1755. No. 1. AIP Publishing LLC, 2016.

[14] Robertson, Leon S., and William Haddon Jr. "The buzzer-light reminder system and safety belt use." American Journal of Public Health 64.8 (1974): 814-815.

[15] Ibrahim, Mochamad, et al. "Buzzer detection and sentiment analysis for predicting presidential election results in a twitter nation." 2015 IEEE international conference on data mining workshop (ICDMW). IEEE, 2015.

A TRACING-BASED TENNIS COACHING AND SMART TRAINING PLATFORM USING ARTIFICIAL INTELLIGENCE AND COMPUTER VISION

Feihong Liu[1], Yu Sun[2]

[1]Crean Lutheran High School, 12500 Sand Canyon Ave, Irvine, CA 92618
[2]California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

ABSTRACT

Athletes in technical sports often find it difficult to analyze their own technique while they're playing [1]. Often, athletes look at the technique of professional players to identify problems they may have. Unfortunately, many types of techniques, such as forehand and backhand swings in tennis, are relatively similar between a beginner and a professional, making it more difficult for comparison. On the other hand, techniques that appear different between professionals and casual can also present different challenges. This is especially true for serves in tennis, where the speed of the swing, the motion of the player, and the angle of the camera recording the player all pose a challenge in analyzing differences between professional and learning tennis players [2]. In this paper, we used two machine learning approaches to compare the serves of two players. In addition, we also developed a website that utilizes these approaches to allow for convenient access and a better experience. We found that our algorithm is effective for comparing two serves of different speeds and synchronized the videos effectively.

KEYWORDS

Pose-estimation, Machine Learning, Scikit-learn.

1. INTRODUCTION

In recent years, Human pose estimation (HPE) has garnered popular attention due to the wide range of its applications in gaming, video surveillance, sign language, etc [3]. Human pose estimation (HPE) is a computer vision technique that aims to accurately recognize the posture of people in its input data (e.g., images, videos, or signals) [4]. HPE accomplishes this through two main steps: localizing anatomical joints, and grouping the joints into a valid configuration [5]. Of course, human pose estimation is complicated by several factors including varying visibility of joints, lighting, and cluttered backgrounds [6]. In fact, these factors challenge even state-of-the-art methods such as Facebook's AI research Detectron2 [7]. Faced with these challenges, many researchers in recent years have started using data taken from multiple dimensions to increase the accuracy of HPE. However, 3D data, although more accurate, was usually much larger than 2D data and took more time to process. Run-time becomes much more important in data with multiple targets. In these cases, either the top-down or bottom-up framework is generally used.

Thanks to its abilities, HPE could be very useful in analyzing individual sports such as tennis, where targets are limited, and motions are predictable. For many players, human pose estimation

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 89-100, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121608

can offer a better way to compare two players' techniques. For techniques such as the serve, it can be difficult to analyze the specific differences because many body parts are moving together. To the observer, the movement of the torso can be misinterpreted as a movement of the arms. Timing can also be an issue, as players tend to serve at varying speeds throughout the serve (having a slow toss of the ball but having a fast swing). Pose estimation offers a way to compare, with more specificity, the techniques of two players.

Improvements in tennis serves are mainly achieved through coaching, either through digital advice, or an in-person coach. Generally, several things are advised for improving a serve: stance, toss, racquet drop, contact spot, and finish [8]. While this seems straightforward, for many casual tennis players, identifying their own issues is the most difficult step. Self-monitored practice is often inaccurate as players themselves misjudge their flaws. Having an in-person coach offers an advantage because coaches are able to make better observations and provide feedback on players' techniques. In addition, coaches are shown to adopt an external focus of attention which can stimulate growth in players [9]. These methods lack a combination of convenience and accuracy. While the advice is straightforward, in-person coaching lacks convenience while self-monitored practice lacks accuracy.

In addition to these methods, other methods also exist to help growth in tennis serves. SwingVision, an editor's choice app backed by various professionals, uses a camera to keep records of tennis points. The app utilizes an iPhone to record points, and detects the types of swing used and other related measurements such as ball speed and the location where it landed. SwingVision is able to detect whether a player used a forehand groundstroke, a slice or an overhead, and even detects whether the player added spin to the ball. This allows for users to see where they can improve on in terms of strategy, and help them learn from their points. Although SwingVision is very useful for reviewing match play, it does not offer ways of improving technique. It gives a way to review the techniques used during a game, but doesn't provide insight into improvement of those techniques. Other services such as TennisGate improves accessibility by making a platform that connects players with online coaches. TennisGate also posts many articles and tutorials on improving technique in tennis. Although this improves on in-person learning by making coaching more convenient, players can often find online coaching less effective and less personalized.

In this paper, we implemented pose estimation to compare two videos of tennis serves. We wanted to make an algorithm that adjusts for different racquet speeds, so a casual tennis player can compare their relatively slow serve with a professional with a faster swing. We also identified keyframes in which players can compare their form with another player. Even with videos that start at different times, this algorithm can identify when the serve starts and edit the video accordingly. We approached doing this through two methods: supervised and unsupervised learning. For both methods, we are using MediaPipe to do pose estimation, which can generate 2d and 3d predictions. For the supervised learning method, we are using various sklearn classification models including Nearest Centroid, Nearest Neighbors analysis and Multiclass SVM. For the unsupervised learning method, we are using SkLearn's k-means clustering to group frames to account for differing speeds [10]

Compared to other approaches, this approach is much more specific. Being specialized in the tennis serve only, this algorithm can give more specific advice on improvements and allow players to form their opinions with a better comparing tool. This algorithm breaks apart the video of tennis players, giving a clear comparison between key moments of their swing and another player's swing and allowing players to compare with more clarity.

The rest of the paper is organized as follows: Section 2 presents related works in our subject; Section 3 gives the details on the challenges that we met during the experiment and designing the sample; Section 4 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 3; Section 5 presents the relevant details about the experiment we did. Finally, Section 6 gives concluding remarks, as well as pointing out the future work of this project.

2. RELATED WORK

Various past research papers discuss similar topics that will occur in this paper. In this section, we will briefly talk about some existing papers published in our subject.

One of the more recent papers, a paper titled *Classification of Tennis Shots with a Neural Network Approach* was published in 2021. This research paper discusses the use of Neural Networks to classify tennis shots [11]. This paper uses data collected from accelerometers, gyroscopes, magnetometers and audio signals to classify tennis shots in five categories: forehand topspin, forehand slice, backhand topspin, backhand slice, and serve. Data is passed into a FCN and a ResNet, and their performance is compared. Our research mainly differs from their research because we are solely focused on classifying various parts of a serve. Our method for obtaining data is purely visual while theirs is biological. In addition, our research is useful in comparison of two serves while their research is useful in the classification of a tennis shot.

In another research paper published by the Lublin University of Technology, human movement analysis was implemented in tennis with a Graph Convolutional Networks approach. This research paper approaches classification of tennis shots through Spatial-Temporal Graph Neural Networks [12]. Data is collected through the Vicon motion capture system, and stored as 3d coordinates and inputted into the neural network to classify between forehand shots, backhands shots and no shots. Our research differs from this research paper as we focus on various parts of the serve while this paper focuses on different swings entirely. This paper also explores the performance of Spatial-Temporal Graph Neural Networks while we explored the different performances of various supervised and unsupervised learning models. In comparison, our research is more applicable and requires little special setup.

Finally, deep learning methods were applied to action recognition for tennis in a research paper titled *Deep Learning for Domain-Specific Action Recognition in Tennis*. This research paper focuses on the performance of certain neural network architectures in application for sports [13]. Visual data from tennis players are collected and features are extracted through a neural network. These features are then passed into another network which is used to classify various tennis actions. This paper also seeks to offer a method to achieve good results for the THETIS dataset, which is composed of low-resolution images of tennis actions. Our research differs from theirs because we seek to synchronize and identify specific parts of tennis serves while theirs seeks to classify different tennis actions. Their paper demonstrated that different neural networks can produce accurate and interpretable results and can potentially be applied to other sports. Our approach provides several methods that a tennis serve can be classified and compares those methods.

3. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

3.1. Adjusting for Varying Starting Points and Speeds for Videos

When comparing two videos, the starting time and speed of the serve may vary. While player 1 can start their swing at the very start of the video, player 2 can start their swing a minute in. In addition, player 1 can also swing a bit slow in the beginning but accelerate their motion while player 2 can maintain a constant speed throughout their serve. This makes comparison of the two videos very complex, as the algorithm ultimately needs to adjust for the speeds and starting points of each player. Initially, we proposed to directly train a neural network to detect when each player is in key stages of a serve. We would train a model with datasets for different stages such as the stance, the racquet drop, and the contact point. However, there were very limited datasets, and would make the algorithm very inefficient. Ultimately, we decided to utilize clustering to adjust for these differences. We would implement clustering first within each video and group frames that are more similar based on the pose of the player. We would then match these clusters from each video with each other to match them up. Doing so, we will generate an array for frames that we have to match up between the videos. Thus, we can then adjust the speeds of the videos between these keyframes to match up key points of the serve and synchronize the two videos.

3.2. Maximizing Accuracy

A common challenge in pose estimation is the different perspectives used for the camera. If two videos of players are taken at a similar perspective, the algorithm performs much better. If videos of players are taken at different perspectives, the algorithm decreases in accuracy. To account for the variety of camera angles, we decided to use an algorithm that uses 3d data to generate results that have relatively little reliance on camera angle. To do this, we took the 3d coordinates of the predicted landmarks and used them to calculate angles between different body parts. We used these angles as data points for both the supervised and unsupervised learning approaches. In addition, we also implemented data augmentation on images for the supervised learning approach to simulate changes in camera angles in real life.

3.3. Data

To classify different stages of a server, a lot of images are required to train an image classification algorithm. However, datasets for images of tennis serves are very scarce. Datasets for tennis mostly concentrated around the historic record for ATP players. Ultimately, we decided that it would be better to create our own dataset. While looking for resources for tennis serves, we found that videos on youtube were generally a good source of tennis serves. We made a program that downloads youtube videos using its url and goes through each frame, allowing the user to manually label each image. We divided the tennis serve into six portions, from the preparation of the serve to the contact point and the finish. To make this process faster, we implemented an algorithm that lets the user look at a specific range of time in a video. After labeling about 20 videos, we got around 400 images for each label. Since a lot of these were quite similar, we manually looked through these images and deleted similar images. We eventually ended up with about 100 images per label. This number is vastly less than what is normally needed for a good algorithm. In addition, these images will undergo a series of algorithms that can eliminate a lot of these images. Thus, we used an algorithm to randomly alter certain features of the image. These augmentations include rotation, skew, zoom and reflection. We generated images with different combinations of these augmentations, and ended up at around 1300 images per label. This allowed the accuracy of our supervised learning models to increase greatly, and also allowed leniency in the input data.

4. SOLUTION

Our algorithm is built in python using the mediapipe and sklearn libraries. To make our algorithm more accessible, we integrated this algorithm using Flask. When certain inputs are given by a user on the website, an HTTP response will be forwarded to the controller, which will allow our algorithm, written in python, to process the data inputted and return values back to the website. Our algorithm does this in several ways, mainly through the two approaches of unsupervised classification and supervised classification.

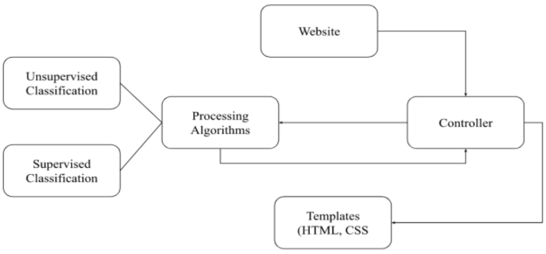


Figure 1. Overview of our solution

4.1. Unsupervised Classification

After a video is uploaded and forwarded to the controller, our algorithm will go through the frames of the videos, identifying a skeleton of the person in the video. Using the angles of the person's joints, we used sklearn's nearest neighbors clustering algorithm to identify five groups of frames within this video. These groups will represent different parts of the serve. After the initial pose-estimation algorithm, an array is generated with elements representing the labels for each frame for each video.

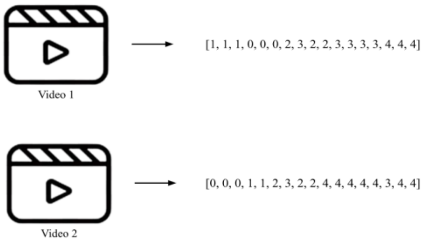
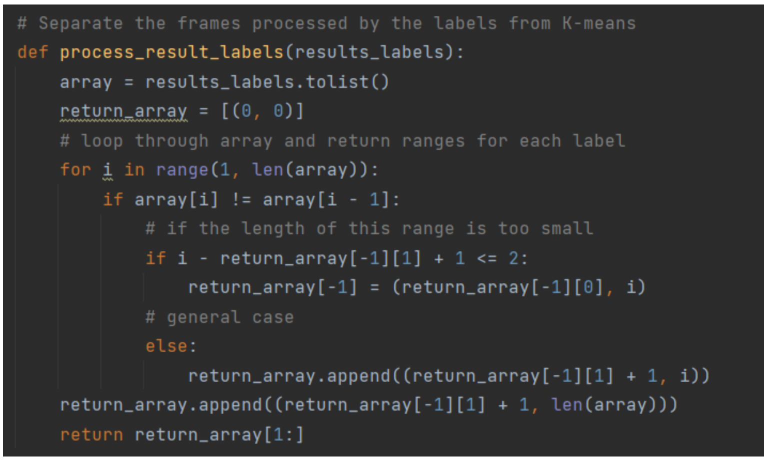


Figure 2. Labels arrays generated for video

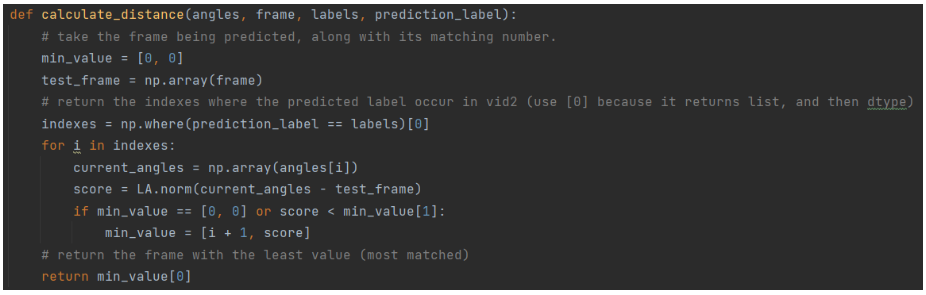
However, these generated frames are not guaranteed to match, so we first implemented an algorithm that can match the groups between the two videos.



```python
# Separate the frames processed by the labels from K-means
def process_result_labels(results_labels):
    array = results_labels.tolist()
    return_array = [(0, 0)]
    # loop through array and return ranges for each label
    for i in range(1, len(array)):
        if array[i] != array[i - 1]:
            # if the length of this range is too small
            if i - return_array[-1][1] + 1 <= 2:
                return_array[-1] = (return_array[-1][0], i)
            # general case
            else:
                return_array.append((return_array[-1][1] + 1, i))
    return_array.append((return_array[-1][1] + 1, len(array)))
    return return_array[1:]
```

Figure 3. Screenshot of code 1

First, the array for video 1 is processed using the algorithm above. A new array is created containing the ranges of the occurrence for the labels. For example, if the input array was [0, 0, 0, 1, 1, 1, 2, 2, 3, 3, 4], the output array will be [(0, 2), (3, 5), (6, 7), (8, 9), (10, 10)] because the 0 occurred at indexes 0 - 2, etc. For simplicity, we added an additional if statement that limits the length of these ranges. If the range is less than 2, then the interval is considered to be too small, and this interval is instead combined with the previous one. In the previous example, instead of outputting [(0, 2), (3, 5), (6, 7), (8, 9), (10, 10)], the algorithm will consider the (6, 7) too small, combining it with (3, 5) to be (3, 7), etc.



```python
def calculate_distance(angles, frame, labels, prediction_label):
    # take the frame being predicted, along with its matching number.
    min_value = [0, 0]
    test_frame = np.array(frame)
    # return the indexes where the predicted label occur in vid2 (use [0] because it returns list, and then dtype)
    indexes = np.where(prediction_label == labels)[0]
    for i in indexes:
        current_angles = np.array(angles[i])
        score = LA.norm(current_angles - test_frame)
        if min_value == [0, 0] or score < min_value[1]:
            min_value = [i + 1, score]
    # return the frame with the least value (most matched)
    return min_value[0]
```

Figure 4. Screenshot of code 2

Using the midpoints of the ranges of the previous arrays, frames of each range in video 1 is compared with frames in video 2. Frames of video 1 are predicted using the model generated from video 2, and then compared with the angles of the frames of video 2 to match certain frames. This process is roughly illustrated in the Figure below.

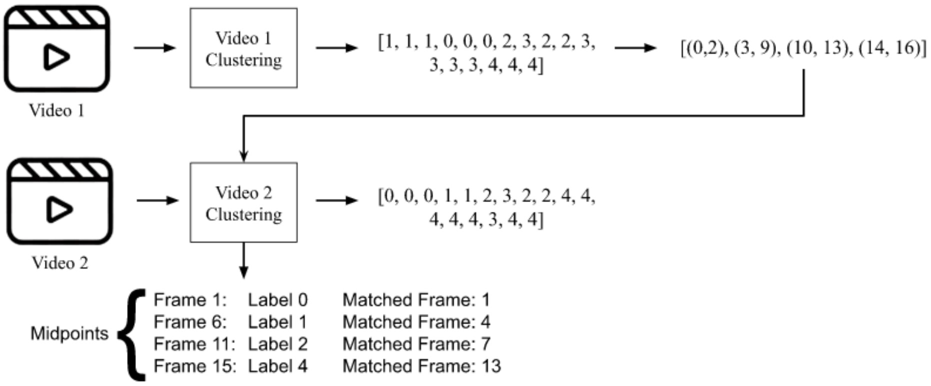


Figure 5. General illustration of our algorithm

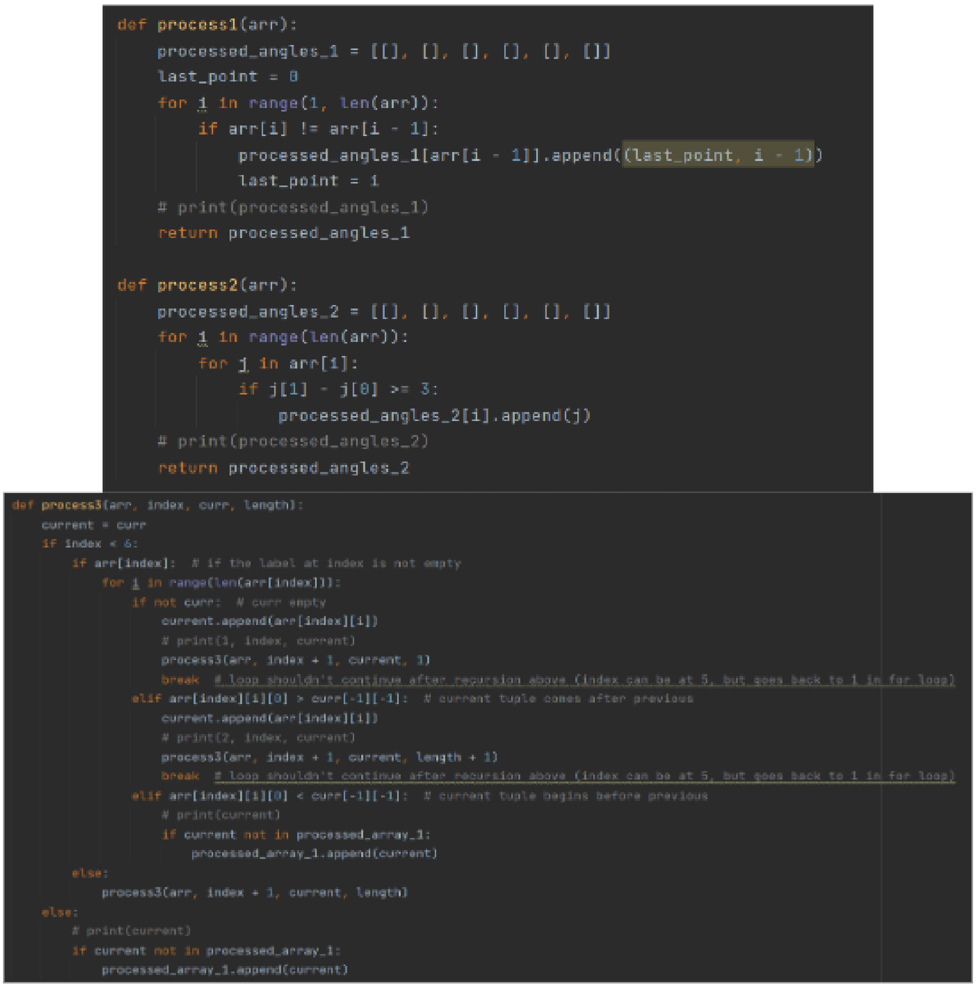
In some cases, however, these matched frames might decrease. Thus, we made an algorithm that selects the longest increasing groups. Afterwards, points where labels change are marked as critical frames, and annotated and saved in a folder. Using the matched frames, we adjusted the speed linearly of the two videos so that the critical frames match in time.

4.2. Supervised Classification

The initial process of the supervised classification approach is similar to the unsupervised approach. Once a user inputs a video, the algorithm will extract angle data from the frames of the video using Mediapipe. However, for supervised classification, a model will be trained using manually collected and labeled images and the angles from the inputted videos will be classified using the trained model.

To train our model, we initially had three approaches: raw images, 3d models or just the angles. Because of the different perspectives for images, the results of the first approach were very inaccurate, while the second approach was impractical as well. We ultimately decided to get the angles for the training images and use those to train the model. We used Nearest Centroid, Nearest Component analysis, Nearest Neighbors and Multiclass SVM as classification models. After a series of experiments to test for the optimal augmentations to apply to images (see experiment 1 and 2), we are ready to implement the model to synchronize videos.

After the angles for the frames of each video are predicted using the trained model, a list is generated with each element representing the predicted label for each frame. However, processing this list proved to be quite different from our approach for unsupervised learning. The main difference is that, for supervised learning, it makes sense only if the predicted labels to be increasing while the unsupervised learning approach might not group frames based on the progression of the serve. We implemented a series of processing algorithms that essentially looks for the longest subsequence where the number of increasing labels are at a maximum (i.e. if a subsequence of length 100 with only label 1 was compared with a subsequence of length 40 but had increasing labels from 1-4, our algorithm would prioritize the latter).



```
def process1(arr):
    processed_angles_1 = [[], [], [], [], [], []]
    last_point = 0
    for i in range(1, len(arr)):
        if arr[i] != arr[i - 1]:
            processed_angles_1[arr[i - 1]].append((last_point, i - 1))
            last_point = i
    # print(processed_angles_1)
    return processed_angles_1

def process2(arr):
    processed_angles_2 = [[], [], [], [], [], []]
    for i in range(len(arr)):
        for j in arr[i]:
            if j[1] - j[0] >= 3:
                processed_angles_2[i].append(j)
    # print(processed_angles_2)
    return processed_angles_2
```

```
def process3(arr, index, curr, length):
    current = curr
    if index < 6:
        if arr[index]:  # if the label at index is not empty
            for i in range(len(arr[index])):
                if not curr:  # curr empty
                    current.append(arr[index][i])
                    # print(1, index, current)
                    process3(arr, index + 1, current, 1)
                    break  # loop shouldn't continue after recursion above (index can be at 5, but goes back to 1 in for loop)
                elif arr[index][i][0] > curr[-1][-1]:  # current tuple comes after previous
                    current.append(arr[index][i])
                    # print(2, index, current)
                    process3(arr, index + 1, current, length + 1)
                    break  # loop shouldn't continue after recursion above (index can be at 5, but goes back to 1 in for loop)
                elif arr[index][i][0] < curr[-1][-1]:  # current tuple begins before previous
                    # print(current)
                    if current not in processed_array_1:
                        processed_array_1.append(current)
        else:
            process3(arr, index + 1, current, length)
    else:
        # print(current)
        if current not in processed_array_1:
            processed_array_1.append(current)
```

Figure 6. Screenshot of processing algorithms

After these processing algorithms are run, we will be able to select particular critical frames and synchronize our videos by linearly adjusting the speed of the videos to match these frames.

5. Experiment

First, we wanted to test the performance of various supervised learning methods in comparison with our initial approach. To do this, we had several algorithms that each used different models for supervised learning. In our experiment, we used Nearest Centroid, Nearest Component Analysis, Nearest Neighbors classification, as well as Multiclass SVM. First, we made an algorithm that allows users to manually label image frames from videos of tennis serves. We iterated through a set of youtube videos and manually labeled them from 1 to 6, each number representing a certain portion of the serve. For example, number 1 represents the preparation for the serve, number 5 represents when the player hits the ball, and other numbers represent other processes throughout the serve. For each label, we collected about 400 images. However, many of these images are very visually similar. After cutting down similar images, we resulted in over 100 images. Because there were very limited methods for obtaining more images of tennis serves, we decided that this number would be enough, but that we can also implement data augmentation to simulate different camera orientations and allow the algorithm to be more accurate with variations in video perspective. The results for each model are shown below.

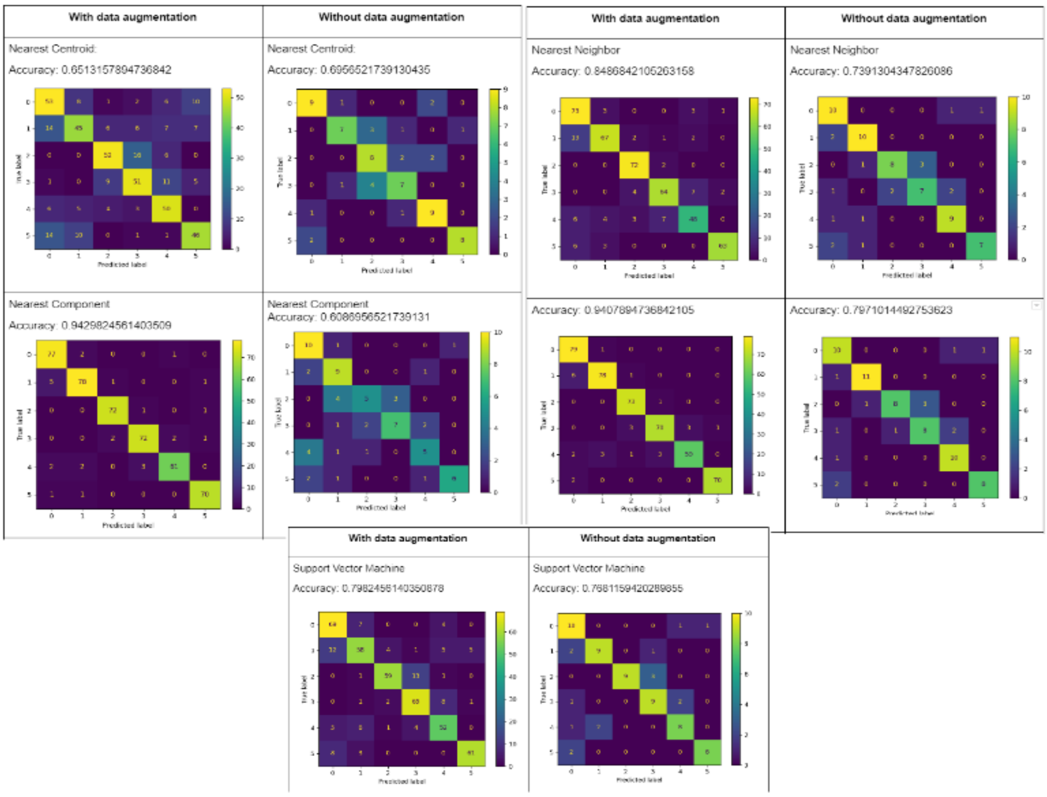


Figure 7. Experiment 1 results (the nearest neighbor approach on the top right uses uniform weights while the one below uses weights dependent on distance)

The addition of data augmentation improved the accuracy of three out of the four supervised learning algorithms. These results are very beneficial to our goal because it demonstrates that supervised learning is a potent tool for our problem, and can be implemented to synchronize videos.

When testing various supervised learning models, we found that certain augmentation features decreased the accuracy of our algorithms while others improved our algorithm. To find what augmentations worked best, we tested different combinations of augmentations and tested the accuracy of each model. Our experiments can be roughly divided into two portions. The first portion of our experiments will test the accuracy of various supervised learning models when using data from augmented images that have one of the four augmentation techniques. Importantly, every image has only one augmentation applied. In the second portion, we will test the accuracy of various supervised learning models using augmented images that have multiple augmentations applied. In this case, every image has at least two augmentations simultaneously applied to it. For both portions, every augmentation will have a 95% chance of acting on the image and 5% chance of doing nothing. This is because we want to preserve a portion of the original images as they are most reliably realistic.

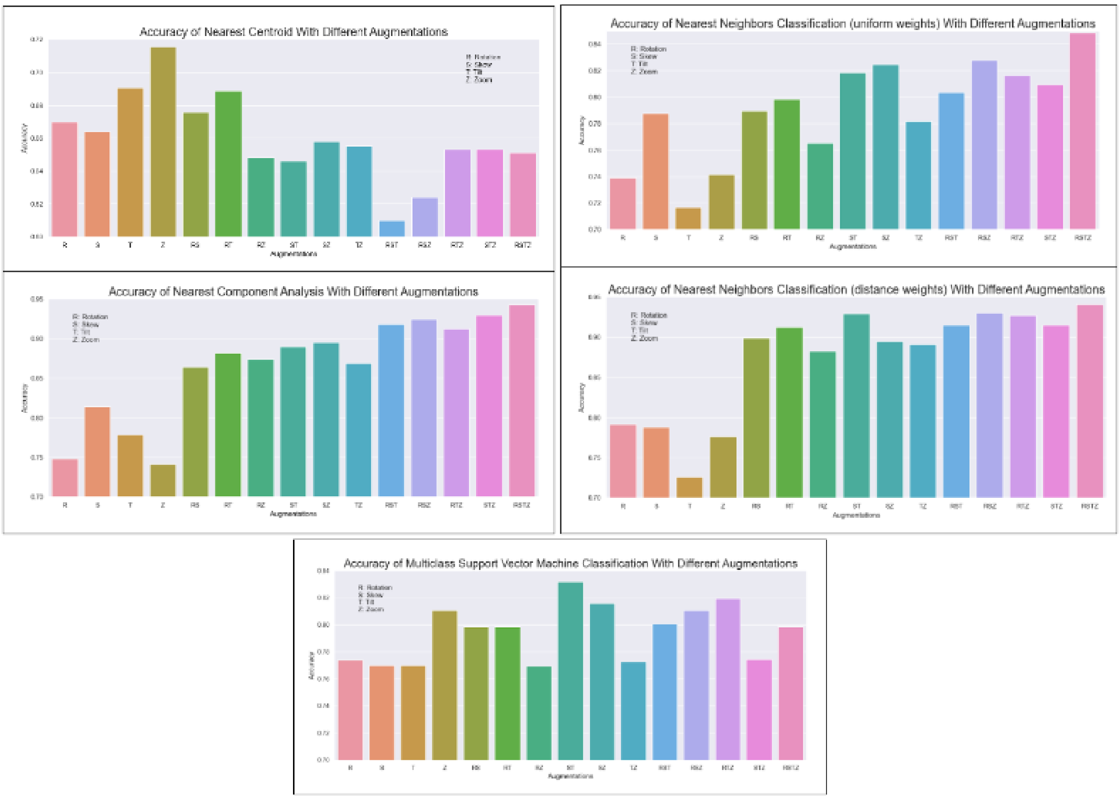


Figure 8. Experiment 2.1 results

The results above show that various models have different performances depending on the augmentations implemented. The Multiclass Support Vector Machine Classifier, for example, performed best when applied to images with Skew and Tilt applied. However, since we want a model that performs best with images of all four augmentations, the Nearest Component Analysis method appears to be the optimal choice.

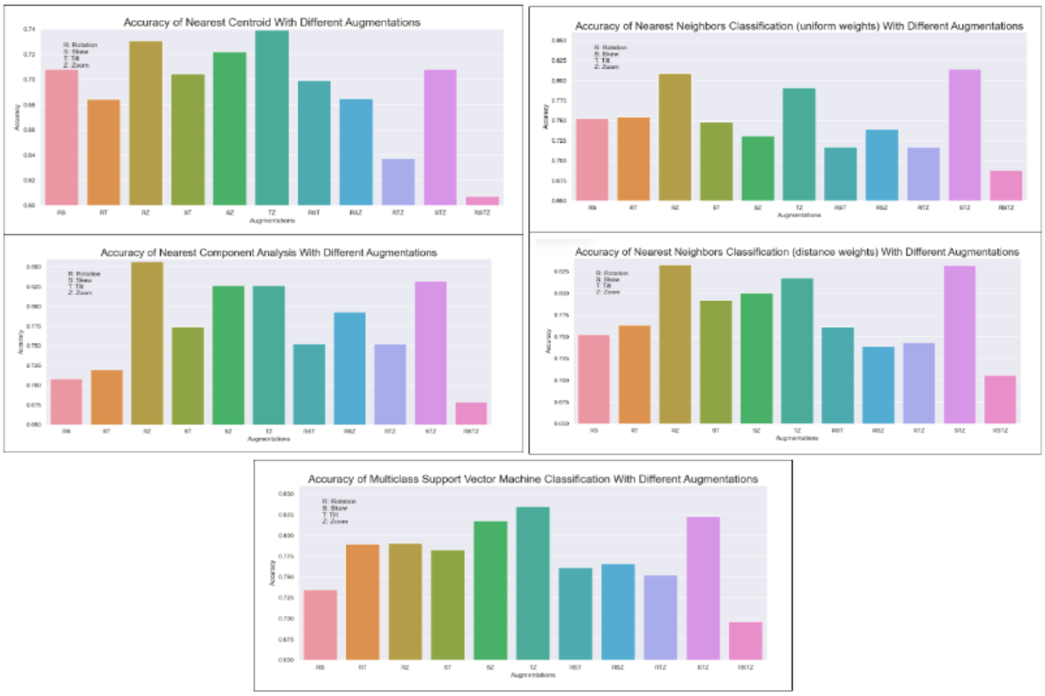


Figure 9. Experiment 2.2 results

The results for portion 2 shows much more fluctuation in accuracy than portion 1. The combination of different augmentations was also generally less effective in improving the accuracy of our model compared to portion 1. However, realistically, our model should utilize both data from portion 1 and portion 2, because camera perspectives are not limited to only one of the listed changes. After combining data from portion 1 and 2, we found that the Nearest Neighbors models with distance weights performed best. Although its accuracy is slightly lower than that of the nearest neighbors component analysis, this model had marginally better runtime.

6. Conclusions

By using various models in supervised and unsupervised learning, we developed an algorithm that offers an easy way to synchronize and compare two serves. We approached this through two ways: supervised and unsupervised learning. For the former, our algorithm will get pose data from all the frames of two uploaded videos. This data will then be used to cluster the frames within each video, respectively, into 6 groups. These groups, one from each video, will then be compared with each other to match similar frames. Using this clustering technique, our algorithm is able to synchronize two videos without manual labeling. The latter was implemented with several models. Data is generated through several algorithms and labeled manually. Using the data given, we used nearest centroid, nearest neighbor, nearest component analysis and multiclass SVM to classify a new frame [14]. This way, two videos can be synchronized as well. This can be very useful for tennis practice, as tennis players can easily compare their serve with another's serve, no matter who it is. Currently, we have created a website using Flask that allows users to use these algorithms and compare their performance. By inputting two videos, the website will give users the option to choose a processing method, and return two videos that are synchronized. Of course, various parts of our algorithm are still flawed in some ways. Firstly, the unsupervised learning approach is very limited by runtime. Videos that are significantly larger will experience significantly longer wait time. While the supervised learning approach performs slightly better with larger video files, the risk of long runtimes is still present. Because the pose data from each frame is taken, the addition of frames will prolong runtime a decent bit. Since the data from videos are compared with each other through various loops in the unsupervised learning approach, the runtime in those processes will increase drastically with file size as well. In addition, our models can be easily fooled by adversarial attacks, because our approach assumes that the video is limited to only the serve. If someone was waving their hand repeatedly in a video, our algorithm might falsely believe that this is a tennis serve and work incorrectly.

Various improvements can be potentially made to improve our algorithm. Getting more data to train our algorithm for supervised learning, for one, can significantly improve the accuracy and consistency of our program. In addition, after identifying the key frames using supervised classification, our processing algorithms sometimes give bad groupings of frames. Thus, improving these processing algorithms can also lead to improvement of the resulting synchronized videos. Finally, training a unique pose-estimation program or modifying our approach can also result in a more efficient algorithm. In the future, many of these aforementioned improvements can be made to offer tennis players a better means of comparing their serves with others.

REFERENCES

[1] Ivanenko, Stanislav, et al. "Analysis of the indicators of athletes at leading sports schools in swimming." Journal of Physical Education and Sport 20.4 (2020): 1721-1726.

[2] Wei, Xinyu, et al. "Predicting serves in tennis using style priors." Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2015.

[3] Andriluka, Mykhaylo, et al. "2d human pose estimation: New benchmark and state of the art analysis." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.

[4] Toshev, Alexander, and Christian Szegedy. "Deeppose: Human pose estimation via deep neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[5] Munea, Tewodros Legesse, et al. "The progress of human pose estimation: a survey and taxonomy of models applied in 2D human pose estimation." IEEE Access 8 (2020): 133330-133348.

[6] Patel, Manisha, and Nilesh Kalani. "A survey on Pose Estimation using Deep Convolutional Neural Networks." IOP Conference Series: Materials Science and Engineering. Vol. 1042. No. 1. IOP Publishing, 2021.

[7] Gao, Fei, et al. "Segmentation-based Background-inference and Small-person Pose Estimation." IEEE Signal Processing Letters (2022).

[8] Myers, Natalie L., et al. "Reliability and validity of a biomechanically based analysis method for the tennis serve." International journal of sports physical therapy 12.3 (2017): 437.

[9] Keller, Martin, Jonas Schweizer, and Markus Gerber. "Pay attention! The influence of coach-, content-, and player-related factors on focus of attention statements during tennis training." European Journal of Sport Science (2022): 1-9.

[10] Lugaresi, Camillo, et al. "Mediapipe: A framework for perceiving and processing reality." Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR). Vol. 2019. 2019.

[11] Ganser, Andreas, Bernhard Hollaus, and Sebastian Stabinger. "Classification of Tennis Shots with a Neural Network Approach." Sensors 21.17 (2021): 5703.

[12] Skublewska-Paszkowska, Maria, Pawel Powroznik, and Edyta Lukasik. "Learning three dimensional tennis shots using graph convolutional networks." Sensors 20.21 (2020): 6094.

[13] Vinyes Mora, Silvia, and William J. Knottenbelt. "Deep learning for domain-specific action recognition in tennis." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017.

[14] Radev, Dragomir R., et al. "Centroid-based summarization of multiple documents." Information Processing & Management 40.6 (2004): 919-938.

FINDING CORRELATION BETWEEN CHRONICAL DISEASES AND FOOD CONSUMPTION FROM 30 YEARS OF SWISS HEALTH DATA LINKED WITH SWISS CONSUMPTION DATA USING FP-GROWTH FOR ASSOCIATION ANALYSIS

Jonas Baschung and Farshideh Einsele

Section of Business Information,
Bern University of Applied Sciences, Switzerland

ABSTRACT

Objective: The objective of the study was to link Swiss food consumption data with demographic data and 30 years of Swiss health data and apply data mining to discover critical food consumption patterns linked with 4 selected chronical diseases like alcohol abuse, blood pressure, cholesterol, and diabetes.

Design: Food consumption databases from a Swiss national survey menu CH were gathered along with data of large surveys of demographics and health data collected over 30 years from Swiss population conducted by Swiss Federal Office of Public Health (FOPH). These databases were integrated and Frequent Pattern Growth (FP-Growth) for the association rule mining was applied to the integrated database.

Results: This study applied data mining algorithm FP-Growth for association rule analysis. 36 association rules for the 4 investigated chronic diseases were found.

Conclusions: FP-Growth was successfully applied to gain promising rules showing food consumption patterns lined with lifestyle diseases and people's demographics such as gender, age group and Body Mass Index (BMI). The rules show that men over 50 years consume more alcohol than women and are more at risk of high blood pressure consequently. Cholesterol and type 2 diabetes is found frequently in people older than 50 years with an unhealthy lifestyle like no exercise, no consumption of vegetables and hot meals and eating irregularly daily. The intake of supplementary food seems not to affect these 4 investigated chronic diseases

KEYWORDS

Data Mining, Association Analysis, Apriori Algorithm, Diet & Chronical Diseases, Health Informatics.

1. INTRODUCTION

Chronical diseases increase in frequency across the globe, becoming an important public health problem even in developing countries. These diseases include obesity, hypertension (blood

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 101-122, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121609

pressure), heart disease, type 2 diabetes, cancer, mental disorders, and many others. They differ from the infectious diseases originated from malnutrition, also called communicable diseases (CD) due to their contagious, dispersive nature. Lifestyle diseases are therefore among the so-called NCD (non-communicable diseases) diseases. According to World Health Organization (WHO), the growing epidemic of chronic diseases afflicting both developed and developing countries are related to dietary and lifestyle changes [1].

Various researchers studied the relationship between nutritional habits and chronic diseases. Schulze et al in [2] discuss current knowledge on the associations between dietary patterns and multiple chronic diseases like cancer, heart disease, stroke, and type 2 diabetes. Their findings confirm that food-based prevention of chronic disease risk should prioritize fruits, vegetables, whole grains, fish and lower consumption of red and processed meats and sugar sweated drinks. Bocedi et al. state in [3] that an unhealthy lifestyle, like unbalanced diet, insufficient sleep, physical inactivity, smoking, alcohol abuse contributes the cause metabolic altercations which can lead to onset of NCDs. Some researchers in [4], [5], [6], [7], [8] studied particularly the impact of the Mediterranean diet, characterized by a high consumption of fruit, vegetables, extra virgin olive oil, cereals, legumes, and fish; a moderate intake of dairy products, eggs, and red wine; and a low intake of animal fats and red meat, as a correct approach to prevent NCDs. Di Marco et al. report in [9] specifically about pasta in Mediterranean diet and its antioxidant compounds like natural bioactive compounds play positive role in the protection of kidney cells from oxidative stress. Koch in [10] demonstrates that an optimal daily intake of antioxidants such as polyphenols and vitamins can counteract the onset of NCDs and to slow their progression. Chen et al. report in [11] that vitamin C (ascorbic acid) and E (tocopherols), are natural compounds that play a pivotal role in preventing the NCDs, mainly for their antioxidant activity. Vitamin C is a water-soluble vitamin, able to protect from the cellular damage exerted by harmful oxidative compounds. Noce et al. report in [12] how ω-3 polyunsaturated fatty acids play a cardioprotective role in male obesity secondary hypogonadism (MOSH) patients. Owen et al. in [13] report of their evaluation of the relationship between dietary quality scores and cardiometabolic risk in a group of older Australian adults, that a high intake of vegetables, grains, and non-processed red meat was associated with a better cardiometabolic risk profile.

Data Mining for chronic diseases prediction and prevention linked with nutritional habits have been explored by different researchers. Lee et al conducted a study using stepwise logistic regression (SLR) analysis, decision tree, random forest, and support vector machine as an alternative and complement to the traditional statistical approaches to identify the factors that affect the health-related quality of life (HRQoL) of the elderly with chronic diseases and to subse-quently develop from such factors a prediction model [14]. D. Qudsi and al. report in [15] from a study that aims to identify the potential benefits that data mining can bring to the health sector, using Indonesian Health Insurance company data as case study. Decision tree as a classification data mining method, was used to generate the prediction model by visualizing the tree to perform predictive analysis of chronic diseases. Z. Lei et al report in [16] of studying the relationship between nutritional ingredients and diseases such as diabetes, hypertension, and heart disease by using data mining methods. They have identified the first two or three nutritional ingredients in food that can benefit the rehabilitation of those diseases. R. McCabe et al. report in [17] of creating a simulation test environment using characteristic models of physician decision strategies and simulated populations of patients with type 2 diabetes, they state of employing a specific data mining technology that predicts encounter-specific errors of omission in representative databases of simulated physician patient encounters and test the predictive technology in an administrative database of real physician-patient encounter data. D.W. Haslam and W.P. James report in [18] of an investigation in a population - based sample of 1140 children performed to derive dietary patterns related to children's obesity status. Their findings reveal that Rules derived through a data mining approach revealed the detrimental influence of the increased consumption of fried food, delicatessen meat, sweets, junk food and soft drinks. K. Lange et al. state in [19] that big data studies may ultimately lead to personalized genotype-based nutrition which could permit the prevention of diet-related diseases and improve medical therapy. A. Hearty and M. Gibney evaluate the usability of supervised data mining methods as ANNs and decision trees to predict an aspect of dietary quality an aspect of dietary quality based on dietary intake with a food-based coding system and a novel meal-based coding system [20]. A. von Reusten et al. used data from 23 531 participants of the EPIC-Potsdam study to analyze the associations between 45 single food groups and risk of major chronic diseases, namely, cardiovascular diseases (CVD), type 2 diabetes and cancer using multivariable-adjusted Cox regression. Their results show that higher intakes of low-fat dairy, butter, red meat, and sauce were associated with higher risks of chronic diseases [21]. E. Yu et al. demonstrate in [22] the usability of supervised data mining methods to extract the food groups related to bladder cancer. Their results show that beverages (non-milk); grains and grain products; vegetables and vegetable products; fats, oils, and their products; meats and meat products were associated with bladder cancer risk.

As a proof of concept, we conducted a preliminary study [23], in which we used a big database gained from a grocery store chain over a certain period along with associated health data of the same region. Association rule mining was successfully used to describe and predict rules linking food consumption patterns with lifestyle diseases. Additionally, we conducted a further study using a medium-sized real-world health and nutritional data from Swiss population and gained interesting rules which showed the link between nutritional habits and chronical diseases. [24] and later another study, in which we used the same national Swiss dietary survey with a five times larger dataset (collected over 25 years) from the national Swiss health survey including demographical information [25]. Based on the finding of the previous studies, where it used the pure Apriori algorithm which resulted that some critical health-related dietary features were pruned out early in course of data mining, we have applied the Weighted Association Mining Rules (WARM) analysis to the latter study.

In This study, we have enlarged our health date from 1991 to 2017 and used Frequent Pattern Growth (FP-Growth) algorithm to gain rules that show the link between Swiss nutritional habits and chronical diseases. Additionally, in this study we added BMI, age group and gender to our candidate patterns, which helped us to gain more specific association rules which contain demographical information when assessing the relationship between chronic diseases and nutrition.

2. Selection, Cleaning, Transformation of the Databases

The following formatting rules must be followed strictly.

2.1. Selection

The data comes from the national surveys menuCH and the health survey that were carried out in Switzerland. The national food survey menuCH (BLV, Federal Office for Food Safety and Veterinary 2020) was carried out for the first time from January 2014 to February 2015. Over 2000 people living in Switzerland were asked about their eating habits and food consumption. The data resulting from the survey is the first representative, national nutritional survey data available in Switzerland from BLV. The second data source comprises health data on the state of health and health-related behavior of the Swiss resident population over a period of 30 years. The Federal Statistical Office (2021) has been collecting health data from the population living in Switzerland every five years using a writ-ten and telephone questionnaire. As part of this study, representative data from around 85,000 people from 1992, 1997, 2002, 2007, 2012 are available. This data has already been pre-cleaned, attributes have been partially selected from the database and the data has been already transformed as reported in [25]. In addition to this, the author has added the health data of 2017 to the health database.

2.2. Cleaning

2.2.1. Cleaning menuCH database

MenuCH database was remained untouched, as described in the study of Mewes et al. [24].

2.2.2. Cleaning health database

From the health data, all records that contained blank or missing survey responses were removed so that only complete data sets are used for analysis. Data cleaning resulted in a significant reduction in the number of usable records for all diseases. For disease alcohol consumption, for example, the number of records was reduced from the original 108,267 to 12,685 records.

For this purpose, all health responses of all persons (always 8 responses per person) including the demographic characteristics were exported from the SQL database for each disease examined, loaded into an Excel, and examined for completeness. The incomplete data sets were eliminated and a new file with complete data sets was then exported from Excel to CSV and used for the investigation in Python. Finally, we have built categories for the 4 selected chronical diseases as follows:

Categories Alcohol:

- 0-17 gr. Alcohol consumption daily,
- 18-22 gr. alcohol consumption daily
- 23-28 gr. Alcohol consumption daily
- more than 28 gr. Alcohol consumption daily

Categories Blood pressure

- not medically assessed normal
- medically judged normal
- not medically judged too low
- medically judged too high
- not medically judged too high
- medically judged too low

Categories Cholesterol

- not medically judged normal
- medically judged normal
- medically judged too high
- not medically assessed too high

Categories Diabetes

- not medically assessed, no diabetes
- medically assessed no diabetes

- medically assessed diabetes
- not medically assessed, diabetes

2.3. Transformation

For the integration of the nutrition and health databases, a third person profile table had to be created, which connects the person profile tables of the nutrition database and the health database. Six attributes were selected which were available in both databases for the personal description:

- Gender (m / f)
- age group (15-29 / 30-39 / 40-49 / 50-64 / 65+)
- Household size (1/2/3/4/5 / 6+)
- Marital status (single / married or registered / widowed / divorced / other)
- Language (de / fr / it).

The selected attributes and their categories resulted in 720 different categories. The PersonIDs in the Menu-CH database and the PersonIDs in the Health database were each assigned to a person category in the PersonProfil table as shown in Fig. 1.

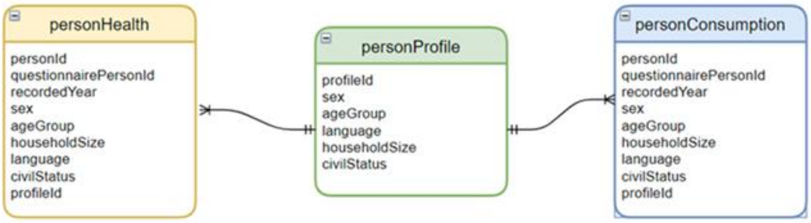


Figure 1. Linking table

Moreover, to work with the 4 selected chronical diseases more efficiently, we decided to add a dimensional scheme to our previously relational database reported in (Lustenberger, 2021) and build 5 dimensions around the fact table personhealth. These dimensions are the 4 selected chronical diseases Alcohol, Blood Pressure, Cholesterol, and diabetes type 2 along with BMI dimension. (see Fig. 2)

2.4. Integration

Fig. 2 shows the scheme of the integrated database with the personProfile as the central link, the structure of the menu-CH data and the connection to health database containing health data from 1991 to 2017, which is a hybrid relational & dimensional scheme containing the relational tables (bottom left) and the dimensional tables (top right). Fig. 2 shows our new hybrid database scheme.

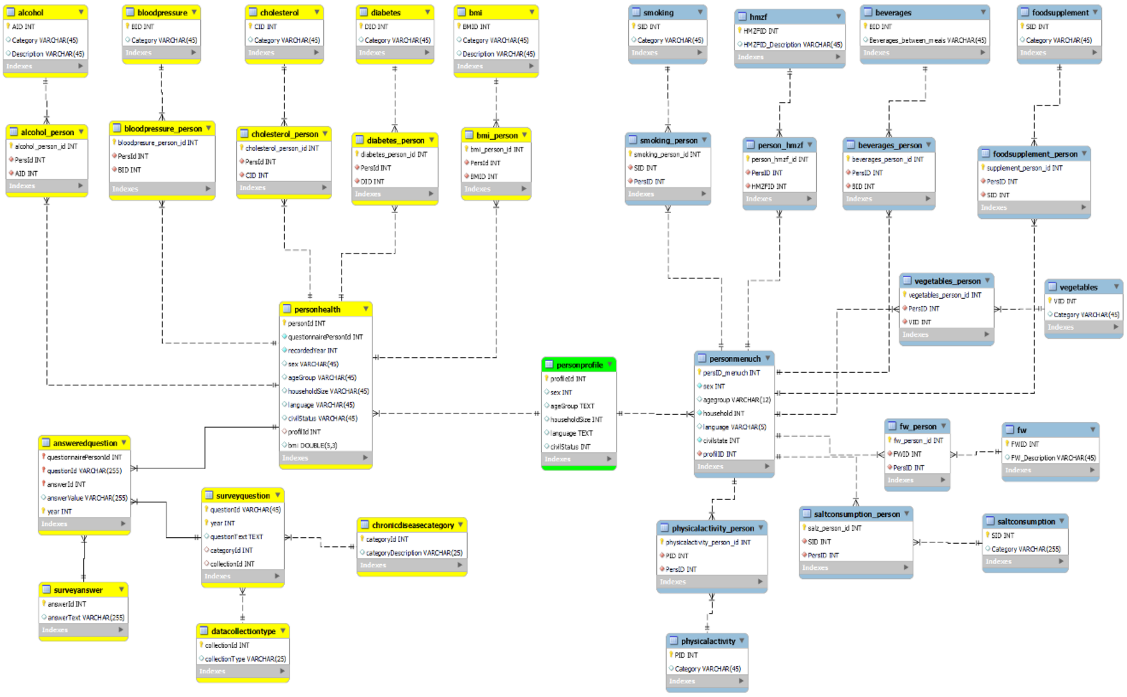
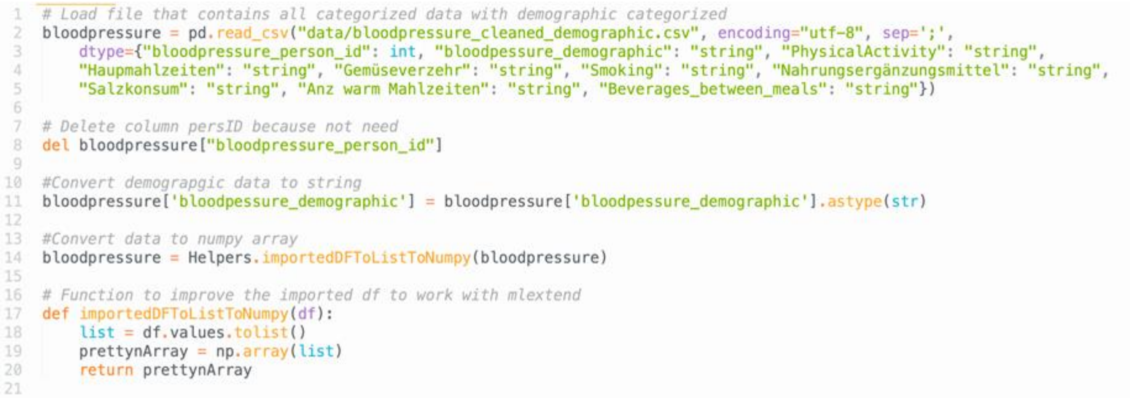


Figure 2. Scheme of the integrated hybrid (relation & dimensional) database

3. ASSOCIATION ANALYSIS WITH FP-GROWTH ALGORITHM

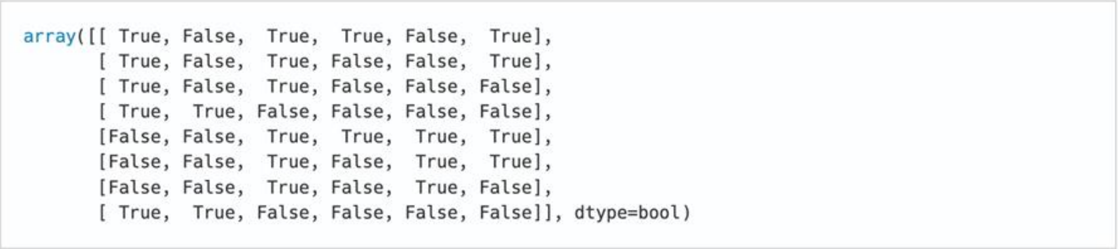
Since our database has grown compared to the previous studies, we decided to use FP-Growth algorithm in this study. FP Growth has been proven to be more performant than Apriori algorithm as a frequent pattern is generated without the need for candidate generation in contrast to Apriori algorithm. FP growth algorithm represents the database in the form of a tree called a frequent pattern tree or FP tree. This tree structure will maintain the association between the itemsets. The database is fragmented using one frequent item. This fragmented part is called “pattern fragment”. The itemsets of these fragmented patterns are analyzed. Thus, with this method, the search for frequent itemsets is reduced accordingly. We have created a total of four Python scripts, in which the analyzes for the four chronic diseases alcohol consumption, high blood pressure, cholesterol and diabetes were carried out. Common functions, which are executed by all four Python scripts, were refactored into a separate "Helper" class. In order to be able to process the data to be analyzed using the methods provided by «mlxtend.frequent_patterns», these should have been available in a one-dimensional, encoded numpy array. The Excel file resulting from the data cleansing was converted into an importable CSV file. The CSV file was loaded directly into the respective Python script and then converted.



```
# Load file that contains all categorized data with demographic categorized
bloodpressure = pd.read_csv("data/bloodpressure_cleaned_demographic.csv", encoding="utf-8", sep=';',
    dtype={"bloodpressure_person_id": int, "bloodpessure_demographic": "string", "PhysicalActivity": "string",
    "Haupmahlzeiten": "string", "Gemüseverzehr": "string", "Smoking": "string", "Nahrungsergänzungsmittel": "string",
    "Salzkonsum": "string", "Anz warm Mahlzeiten": "string", "Beverages_between_meals": "string"})

# Delete column persID because not need
del bloodpressure["bloodpressure_person_id"]

#Convert demograpgic data to string
bloodpressure['bloodpessure_demographic'] = bloodpressure['bloodpessure_demographic'].astype(str)

#Convert data to numpy array
bloodpressure = Helpers.importedDFToListToNumpy(bloodpressure)

# Function to improve the imported df to work with mlextend
def importedDFToListToNumpy(df):
    list = df.values.tolist()
    prettynArray = np.array(list)
    return prettynArray
```

Figure 3. Python script for preparation of data

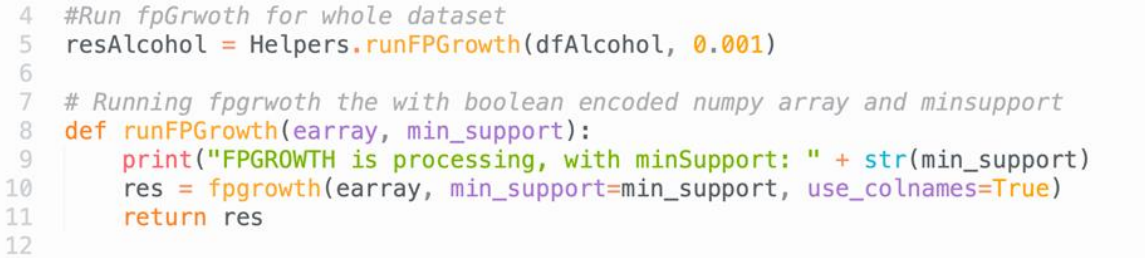
Converting it to a one-dimensional numpy array required several steps. The characteristics sometimes contain longer answers separated by commas, such as "Water, coffee/tea, SFGEI, Light, Milchg." when consuming drinks. This caused Numpy to convert the imported dataframe into a multidimensional array. For this reason, a function was created which successfully converts the dataframe into a numpy array. In a next step, the existing array is converted from a "pandas" dataframe into a boolean numpy array, which can be used directly to search for frequent item sets.



```
array([[ True, False,  True,  True, False,  True],
       [ True, False,  True, False, False,  True],
       [ True, False,  True, False, False, False],
       [ True,  True, False, False, False, False],
       [False, False,  True,  True,  True,  True],
       [False, False,  True, False,  True,  True],
       [False, False,  True, False,  True, False],
       [ True,  True, False, False, False, False]], dtype=bool)
```

Figure 4. Numpy array

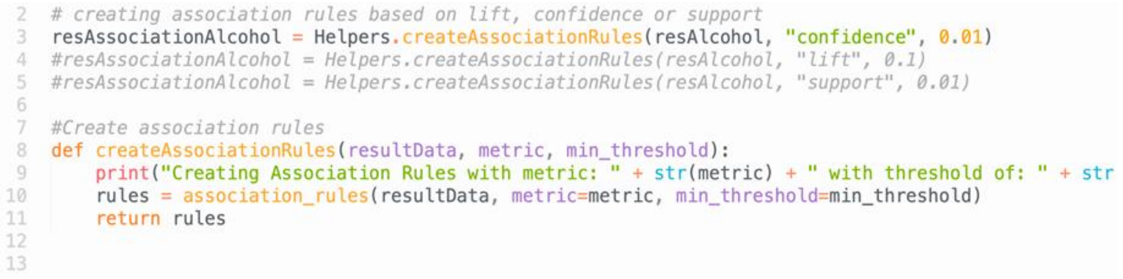
The occurrence of each feature in each transaction is stored in this array, which makes it easier to find frequent item sets. The Boolean data frame that is now available can be passed to the FP-Growth function (Fig. 6) to generate frequent item sets.



```
#Run fpGrwoth for whole dataset
resAlcohol = Helpers.runFPGrowth(dfAlcohol, 0.001)

# Running fpgrwoth the with boolean encoded numpy array and minsupport
def runFPGrowth(earray, min_support):
    print("FPGROWTH is processing, with minSupport: " + str(min_support))
    res = fpgrowth(earray, min_support=min_support, use_colnames=True)
    return res
```

Figure 5. Python script to run the data frame

The execution of the FP growth algorithm was carried out in several iterative runs for each chronic disease, in which the results are examined with different support values. It was started with a support of 1%. Since the number of sick people is small in relation to the healthy people in the existing data set, disease characteristics appear less often in the results. Association rules could be set up from the frequently generated item sets. For this purpose, the Python code was implemented and tested. Afterwards, association rules were generated iteratively for each of the four chronic diseases alcohol consumption, hypertension, cholesterol, and diabetes. The rules were first created based on the parameters support, confidence, and lift, then filtered, sorted and exported. Once exported, they were further examined, sorted, and filtered in Excel.



```python
# creating association rules based on lift, confidence or support
resAssociationAlcohol = Helpers.createAssociationRules(resAlcohol, "confidence", 0.01)
#resAssociationAlcohol = Helpers.createAssociationRules(resAlcohol, "lift", 0.1)
#resAssociationAlcohol = Helpers.createAssociationRules(resAlcohol, "support", 0.01)

#Create association rules
def createAssociationRules(resultData, metric, min_threshold):
    print("Creating Association Rules with metric: " + str(metric) + " with threshold of: " + str
    rules = association_rules(resultData, metric=metric, min_threshold=min_threshold)
    return rules


```

Figure 6. Python script to create association rules using FP-Growth algorithm

Fig. 6 shows how wanted and unwanted conclusions are stored in the variables search for Consequents and not needed Consequents (neither of which are shown completely). These were filtered out of the data frame with the generated association rules using regular expressions. After filtering, the association rules were each exported in a single CSV file for the parameters support, confidence, and lift. The CSV files could then be imported into Excel files and the data could be further analyzed there.

4. Gained Association Rules

A total of 36 new association rules, 10 rules for alcohol abuse, 9 rules for blood pressure (hypertension), 9 rules for cholesterol and 8 rules for diabetes containing the relevant diet and the newly added demographics such as gender, age group and BMI of Swiss residents could be found. We have grouped them based on their highest lift, support, and confidence. The highest lift shows xxx, the highest support shows xxx and the highest confidence shows xxx.

4.1. Association rules for alcohol abuse

4.1.1. Association rules with the highest lift

Rule 1

1.2% of the sample consumed more than 28 gr of alcohol daily, were men, older than 65 years and had a BMI between 18.5-24.9 and had the following characteristics:

- They rarely move (0 days a week)
- They use salt without additives, without re-salting at home
- They regularly eat breakfast, lunch, and dinner

This rule meets in 4.1% of the sample (confidence) and has a lift of 3.5

Rule 2

1.2% of the sample consumed more than 28 gr of alcohol daily, were men, older than 65 years and had a BMI between 18.5-24.9 and had the following characteristics:

- They rarely move (0 days a week)
- They use salt without additives, without re-salting at home
- They regularly eat breakfast, lunch, and dinner
- They never to rarely eat a hot meal

This rule is true in 2% of the sample (confidence) and has a lift of 1.6

Rule 3

1.5% of the sample consume daily 0-17 gr. Alcohol, are women, between 15-29 years old with a BMI below 18.5 and have the following characteristics:

- They regularly exercise between five to seven times a week
- They never to rarely eat a hot meal
- They never to rarely (0-3x per month) eat vegetables
- They use salt without additives, without re-salting at home
- They drink water, coffee, or tea between meals
- They regularly eat breakfast, lunch, and dinner

This rule is true in 70% of the sample (confidence) and has a lift of 63

Rule 4

1.5% of the sample consume daily 0-17 gr. Alcohol, are women, between 50-64 years old with a BMI below 18.5 and have the following characteristics:

- They regularly exercise between five to seven times a week
- They never to rarely eat a hot meal
- They never to rarely (0-3x per month) eat vegetables
- They use salt without additives, without re-salting at home
- They drink water, coffee, or tea between meals
- They regularly eat breakfast, lunch, and dinner

This rule is true in 59% of the sample (confidence) and has a lift of 44

Rule 5

1.3% of the sample consume daily 0-17 gr. Alcohol, are men, between 50-64 years old with a BMI below 18.5 and have the following characteristics:

- They rarely move (0 days a week)
- They never to rarely eat a hot meal
- They never to rarely (0-3x per month) eat vegetables
- They use salt without additives, without re-salting at home
- They drink SFGEI (assumption: sugar-free beverages) between meals.
- They regularly eat breakfast, lunch, and dinner

This rule is true in 51% of the sample (confidence) and has a lift of 36

4.1.2. Association rules with the highest support

Rule 6

7.3% of the sample consume daily 0-17 gr. Alcohol, are women, between 50-64 years old with a BMI 18.5-24.9 and have the following characteristics:

- They never to rarely eat a hot meal
- They never to rarely (0-3x per month) eat vegetables
- They use salt without additives, without re-salting at home
- They regularly eat breakfast, lunch and dinner
- They move irregularly (1-4 days)

This rule is true in 14% of the sample (confidence) and has a lift of 1.92

Rule 7

7.3% of the sample consume daily 0-17 gr. Alcohol, are women, between 50-64 years old with a BMI 18.5-24.9 and have the following characteristics:

- They never to rarely eat a hot meal
- They never to rarely (0-3x per month) eat vegetables
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner

This rule is true in 9.4% of the sample (confidence) and has a lift of 1.28

4.1.3. Association rules with the highest confidence

Rule 8

4.89% of the sample consume daily 0-17 gr. Alcohol, are women, 65 years and older with BMI 25-29.9 and have the following characteristics:

- They never to rarely eat a hot meal
- They never eat vegetables irregularly (1-2x week)
- They do not take dietary supplements
- They use salt without additives, without re-salting at home
- They move irregularly (1-4 days per week)
- They regularly eat breakfast, lunch, and dinner
- They drink coffee/tea and milk drinks (between meals)

This rule is true in 100% of the sample (confidence) and has a lift of 20

Rule 9

7.34% of the sample consume daily 0-17 gr. Alcohol, are women, between 50-64 years old with a BMI 18.5-24.9 and have the following characteristics:

- They never to rarely eat a hot meal
- They never to rarely eat vegetables
- They regularly eat breakfast, lunch, and dinner

- They drink coffee/tea and SFGEI (between meals).
- They do not take dietary supplements
- They use salt without additives, without re-salting at home
- They move irregularly (1-4 days per week)

This rule is true in 61% of the sample (confidence) and has a lift of 8.4

Rule 10

Rules with high alcohol consumption have a lower confidence because they occur less often, nevertheless two rules with the highest confidence from this category were established:

1.19% of the sample consumed more than 28 gr of alcohol daily, were men, 65 years and older, and had a BMI between 18.5-24.9 and had the following characteristics:

- They never to rarely eat a hot meal
- They never to rarely eat vegetables
- They regularly eat breakfast, lunch, and dinner
- They use salt without additives, without re-salting at home

This rule meets in 1.95% of the sample (confidence) and has a lift of 1.6.

4.2. Association rules for blood pressure (hypertension)

In our health database, we had records with medically assessed and medically not assessed hypertension. Table 1 shows this distribution in the database. However, we applied data mining on records which were medically assessed.

Table 1. Distribution of medically assessed records for blood pressure

Expression	**Num. of transactions**	**Share in %**
not medically assessed normal	28'889	65,69
medically assessed normal	6'662	15,15
not medically assessed too low	4'942	0,35
medically judged too high	1'812	4,12
not medically assessed too high	1'520	3,46
medically judged too low	154	11,24
Total	43'979	100

4.2.1. Association rules with highest lift

Rule 1

1.27% of the sample have medically judged normal blood pressure, are women, 65 years and older, have a BMI of 25-29.9, and have the following characteristics:

- They move irregularly (1-2x / week)
- They use salt without additives, without re-salting at home
- They do not take dietary supplements
- They drink coffee/tea, milk drinks (between meals).

- They regularly eat breakfast, lunch, and dinner
- They never to rarely eat hot meals
- They do not smoke and did not smoke before
- They prepare irregularly (1-2x / week)) vegetables

This rule is true in 32% of the sample (confidence) with a lift of 17.8

Rule 2

0.23% of the sample have medically assessed hypertension, are women, 65 years and older, have a BMI of 18.5 - 24.9, and have the following characteristics:

- They do not take dietary supplements
- They use salt without additives, without re-salting at home
- They never to rarely eat vegetables irregularly (0-3x / month)
- They drink coffee/tea (between meals)
- They eat hot meals irregularly (4-7x / week)
- They exercise regularly (5-7 days per week)
- They do not smoke and did not smoke before
- They regularly eat breakfast, lunch, and dinner

This rule is true (confidence) in 9% of the sample with a lift of 22.3

Rule 3

0.13% of the sample have medically assessed hypertension, are men, 65 years and older, have a BMI of >= 30, and have the following characteristics:

- They do not take dietary supplements
- They use salt without additives, without re-salting at home
- They never to rarely eat vegetables irregularly (0-3x / month)
- They drink coffee/tea
- They regularly eat breakfast, lunch, and dinner
- They never to rarely eat hot meals (0-3x / week)

This rule is true in 2.6% of the sample (confidence) with a lift of 19.

4.2.2. Association rules with highest support

Rule 4

0.7% of the sample have medically assessed hypertension, are men, 65 years and older, have a BMI of 18.5-24.9, and have the following characteristics:

- They do not take dietary supplements
- They use salt without additives, without re-salting at home
- They never to rarely eat vegetables (0-3x / month)
- They drink coffee/tea (between meals)
- They regularly eat breakfast, lunch, and dinner
- They never to rarely eat hot meals (0-3x / week)

- They do not smoke and have not smoked before

This rule is true in 1.2% of the sample (confidence) with a lift of 1.7

Rule 5

0.33% of the sample have medically assessed hypertension, are women, 65 years and older, have a BMI of 25-29.9, and have the following characteristics:

- They do not take dietary supplements
- They use salt without additives, without re-salting at home
- They eat vegetables irregularly (1-2x / week)
- They move irregularly (1-4 days / week)
- They drink coffee/tea and milk drinks (between meals)
- They regularly eat breakfast, lunch, and dinner
- They never to rarely eat hot meals (0-3x / week)
- They do not smoke and did not smoke before

This rule is true in 8.4% of the sample (confidence) with a lift of 17

Rule 6

2.74% of the sample have medically assessed normal blood pressure, are men, 65 years and older, have a BMI of 24.9-29.9, and have the following characteristics:

- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They never to rarely eat hot meals
- They never - rarely eat vegetables (0-3x month)

This rule is true in 4.5% of the sample (confidence) with a lift of 1.5

4.2.3. Association rules with highest confidence

Rule 7

2.9% of the sample have medically assessed normal blood pressure, are men, 65 years and older, have a BMI of 24.9-29.9, and have the following characteristics:

- They drink alcoholic beverages
- They rarely to never move
- They rarely - never eat vegetables
- They do not take dietary supplements

This rule is true in 36% of the sample (confidence) with a lift of 12

Rule 8

0.41% of the sample have medically assessed high blood pressure, are women, 65 years and older, have a BMI of 18.5-24.9, and have the following characteristics:

- They eat vegetables irregularly (1-2x / week)
- They eat hot meals irregularly (4-7x/week)
- They move regularly (5-7 days / week)
- They use salt without additives, without re-salting at home

This rule is true in 9.2% of the sample (confidence) with a lift of 22

Rule 9

0.41% of the sample have medically assessed high blood pressure, are women, 65 years and older, have a BMI of 18.5-24.9, and have the following characteristics:

- They do not take dietary supplements
- They use salt without additives, without re-salting at home
- They eat vegetables irregularly (1-2x / week)
- They move irregularly (1-2x / week)
- They drink coffee/tea and milk drinks (between meals)
- They never to rarely eat hot meals (0-3/month)
- They do not smoke and did not smoke before

This rule is true in 8.4% of the sample (confidence) with a lift of 17

4.3. Association rules for cholesterol

In our health database, we had records with medically assessed and medically not assessed cholesterol. Table 2 shows this distribution in the database. However, we applied data mining on records which were medically assessed.

Table 2. Distribution of medically assessed records for cholesterol

Expression	**No. of transactions**	**Share in %**
not medically assessed normal	26'812	80,57
medically assessed normal	3'936	11,83
medically judged too high	1'364	4,10
not medically assessed too high	1'164	3,50
Total	33'276	100

4.3.1. Association rules with highest lift

Rule 1

0.28% of the sample had medically assessed normal cholesterol, were men, 65 years and older, had a BMI of 25-29.9, and had the following characteristics:

- They drink water and alcoholic beverages
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They do not smoke and have not smoked before
- They rarely to never eat vegetables (0-3x/month)
- They rarely to never eat hot meals (0-3x / week)
- They use salt without additives, and regular re-salting at home (1-5/10 meals).

This rule is true in 0.15% of the sample (confidence) with a lift of 108

Rule 2

0.27% of the sample had medically assessed high cholesterol, were women, 65 years and older, had a BMI of 18.5-24.9, and had the following characteristics:

- They use salt without additives, without regularly re-salting at home
- They drink water, tea/coffee (between meals)
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They eat hot meals irregularly (4-7x / week)
- They eat vegetables irregularly (1-2x/week)

This rule is true in 4.7% of the sample (confidence) with a lift of 17

Rule 3

0.13% of the sample had medically assessed high cholesterol, were women, 65 years and older, had a BMI of >=30, and had the following characteristics:

- They use salt without additives, without regularly re-salting at home
- They drink water, tea/coffee (between meals)
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They rarely to never eat hot meals (0-3x / week)
- They eat vegetables regularly (>2 /week)
- They do not smoke and did not smoke before

This rule is true in 11% of the sample (confidence) with a lift of 79

Rule 4

0.13% of the sample had medically assessed high cholesterol, were men, 65 years and older, had a BMI of >=30, and had the following characteristics:

- They never to rarely eat vegetables (0-3 / month)
- They use salt without additives, without regularly re-salting at home
- They rarely to never eat hot meals (0-3x / week)
- They drink coffee/tea (between meals)

This rule is true in 2.4% of the sample (confidence) with a lift of 17

4.3.2. Association rules with highest support

Rule 5

2.4% of the sample had medically assessed normal cholesterol, were men, 65 years and older, had a BMI of 18.5-24.9, and had the following characteristics:

- They use salt without additives, without regularly re-salting at home
- They do not take dietary supplements

- They regularly eat breakfast, lunch, and dinner
- They do not smoke and did not smoke before
- They rarely to never eat vegetables (0-3x/month)
- They rarely to never eat hot meals (0-3x / week)

This rule is true in 4.1% of the sample (confidence) with a lift of 1.7

4.3.3. Association rules with highest confidence

Rule 6

2.4% of the sample had medically assessed normal cholesterol, were women, 65 years and older, had a BMI of 25-29.9, and had the following characteristics:

- They use salt without additives, without regularly re-salting at home
- They move regularly (5-7 days / week)
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They do not smoke and did not smoke before
- They rarely to never eat hot meals (0-3x / week)
- They eat vegetables irregularly (1-2x/week)
- They drink tea/coffee and milk drinks (between meals)

This rule is true in 20.9% of the sample (confidence) with a lift of 15

Rule 7

0.11% of the sample had medically assessed high cholesterol, were men, 65 years and older, had a BMI of 18.5-24.9, and had the following characteristics:

- They use salt without additives, without regularly re-salting at home
- They move regularly (5-7 days / week)
- They never to rarely eat vegetables (0-3x/month)
- They rarely to never eat hot meals (0-3x / week)
- They drink alcoholic beverages (between meals)

This rule is true in 6.06% of the sample (confidence) with a lift of 11

Rule 8

0.27% of the sample had medically assessed high cholesterol, were women, 65 years and older, had a BMI of 18.5-24.9, and had the following characteristics:

- They use salt without additives, without regularly re-salting at home
- They drink water, coffee/tea (between meals)
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They eat hot meals irregularly (4-7x / week)
- They do not smoke and did not smoke before
- They eat vegetables irregularly (1-2x / week)

This rule is true in 4.7% of the sample (confidence) with a lift of 17

4.4. Association rules for diabetes

In our health database, we had records with medically assessed and medically not assessed diabetes. Table 3 shows this distribution in the database. However, we applied data mining on records which were medically assessed.

Table 3. Distribution of medically assessed records for diabetes

Expression	**No. of transactions**	**Share in %**
not medically assessed, no diabetes	26'648	94,08
medically assessed no diabetes	1174	4,14
medically assessed diabetes	328	1,16
not med. assessed, diabetes	174	0,61
Total	28'324	100

4.4.1. Association rules with highest lift

Rule 1

0.21% of the sample have medically assessed diabetes, are men, 65 years and older, have a BMI of 18.5-24.9, and have the following characteristics:

- They never to rarely eat vegetables (0-3x / month)
- They never to rarely eat hot meals (0-3x / week)
- They do not take dietary supplements
- They drink alcoholic beverages (between meals)
- They consume salt without addition never re-salt at home

This rule is true in 1.9% of the sample (confidence) and a lift of 8.9

Rule 2

0.15% of the sample have medically assessed diabetes, are men, 50 to 64 years old, have a BMI of 18.5-24.9, and have the following characteristics:

- They never to rarely eat vegetables (0-3x / month)
- They never to rarely eat hot meals (0-3x / week)
- They do not take dietary supplements
- They drink alcoholic beverages (between meals)
- They consume salt without addition never re-salt at home
- They do not smoke and did not smoke before

This rule is true in 0.21% of the sample (confidence) and a lift of 1.4

Rule 3

0.26% of the sample did not have diabetes as medically assessed, were women, 65 years and older, had a BMI of 18.5-24.9, and had the following characteristics:

- They do not take dietary supplements
- They eat vegetables irregularly (1-2x / week)

- They move regularly (5-7 days / week)
- They consume salt without addition never re-salt at home
- They eat hot meals irregularly (4-7x / week)

This rule is true in 5.7% of the sample (confidence) and a lift of 21

Rule 4

- 0.21% of the sample did not have diabetes as medically assessed, were men, 65 years and older, had a BMI of >= 30 and had the following characteristics:
- They never to rarely eat vegetables (0-3x / month)
- They move irregularly (1-4 days / week)
- They never to rarely eat hot meals (0-3x / week)
- They consume salt without addition never re-salt at home
- They consume coffee/tea (between meals)

This rule is true in 3.9% of the sample (confidence) and a lift of 18.7

4.4.2. Association rules with highest support

Rule 5

0.21% of the sample have medically assessed diabetes, are men, 65 years and older, have a BMI of 18.5-24.9, and have the following characteristics:

- They never to rarely eat vegetables (0-3x / month)
- They never to rarely eat hot meals (0-3x / week)
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They do not smoke and did not smoke before

This rule is true in 0.29% of the sample (confidence) and a lift of 1.4

Rule 6

1% of the sample are medically assessed not to have diabetes, are men, 65 years and older, have a BMI of 18.5-24.9, and have the following characteristics:

- They never to rarely eat vegetables (0-3x / month)
- They never to rarely eat hot meals (0-3x / week)
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They do not smoke and did not smoke before
- They consume salt without addition never re-salt at home

This rule is true in 1.6% of the sample (confidence) and a lift of 1.6

4.4.3. Association rules with the highest confidence

Rule 7

0.22% of the sample have medically assessed diabetes, are men, 65 years and older, have a BMI of 18.5-24.9, and have the following characteristics:

- They never to rarely eat vegetables (0-3x / month)
- They never to rarely eat hot meals (0-3x / week)
- They do not take dietary supplements
- They drink alcoholic beverages (between meals)

This rule is true in 1.9% of the sample (confidence) and a lift of 8.9

Rule 8

0.11% of the sample have medically assessed diabetes, are women, 65 years and older, have a BMI of 25-29.9, and have the following characteristics:

- They consume salt without addition never re-salt at home
- They do not take dietary supplements
- They regularly eat breakfast, lunch, and dinner
- They do not smoke and did not smoke before

This rule is true in 0.14% of the sample (confidence) and a lift of 1.2

Rule 9

0.51% of the sample did not have diabetes as medically assessed, were women, 65 years and older, had a BMI of 25.0-29.9, and had the following characteristics:

- They do not take dietary supplements
- They drink coffee/tea and milk drinks (between meals)
- They do not smoke and did not smoke before
- They never to rarely eat hot meals (0-3x / week)
- They consume salt without addition never re-salt at home
- They regularly eat breakfast, lunch, and dinner

This rule is true in 6.7% of the sample (confidence) and a lift of 12

5. CONCLUSION AND FUTURE WORK

In this study, we applied FP-Growth to find association rules that demonstrate the relationship between nutritional habits and four chronic diseases such as alcohol abuse, blood pressure, cholesterol, and diabetes along with the corresponding demographics. Our data base includes health data from multiple surveys over 30 years (1992-2017) from tens of thousands of Swiss population and nutrition data from the first Swiss nationwide nutritional survey (2014-2015). In the previous studies (Mewes, Einsele, 2020) and (Lustenberger, Einsele, 2021), we have dealt with smaller health databases and applied Apriori algorithm to extract association rules. The nutritional data in the previous study and the current study remains unchanged. Since the health database has grown significantly in this study, we chose FP-Growth and used a Python's machine learning library to be more performant and efficient comparing to Apriori algorithm in the previous studies. Further improvement in the current study is that we have added demographic information about gender, age group and Body Mass Index (BMI) to the list of itemsets to gain more accurate association rules from our integrated database.

The study shows that concerning alcohol abuse men over 65 years that have a normal BMI, who overconsume alcohol daily, rarely move and rarely eat hot meals. Furthermore, middle aged men between 54 and 65 years with normal BMI, who do not overconsume alcohol, have a similar lifestyle to the previous group like making no exercise and rarely eating hot meals and vegetables. Women, on the contrary, who are younger than 64 years old with an ideal BMI don't overconsume alcohol. These women exercise regularly but rarely eat hot meal or vegetables but eat 3 times a day. Hence, although their lifestyle is similar to the men of same age, women tend to less over consume alcohol.

In addition to this, women older than 65 years with normal to high BMI with hypertension, do not smoke, exercise weekly, eat regularly but rarely eat hot meals or vegetables. Men older than 65 with a normal and high BMI with hypertension eat regularly but rarely hot meals or vegetables as well. Additionally, our rules show that intake of dietary supplements does not reduce blood pressure.

Furthermore, women older than 65 years with a normal BMI who have high cholesterol, eat hot meals and vegetables irregularly. Additionally, women over 65 years with a high BMI, eat vegetables both regularly and irregularly, don't smoke, eat rarely hot meals but 3 meals a day. Finally, Men over 65 years old with a normal BMI who have high cholesterol, rarely eat vegetables, and hot meals but drink daily alcoholic beverages. This is in accordance with our found rules about alcohol consume, which shows that older men tend to more overconsume alcohol, and this could result to high cholesterol as well.

Finally, our rules show that mostly men and women over 50 years show diabetes type 2. Which can be a result of years of unhealthy lifestyles. According to our found rules, men over 50 years with normal to high BMI, rarely eat vegetables and hot meals and consume alcoholic beverages daily. Interestingly additional rules show that the same age group no matter which gender who eat three meals a day but eat rarely hot meals and vegetables have no diabetes. This could be an important hint that eating three meals a day by people older than 50 could reduce the risk of diabetes 2. For future work, it is essential to expand the data base. The data base is unevenly distributed with two different data sources menuCH and SGB (health data). With 2'000 persons from the nutrition survey (menuCH) and a total of about 120'000 persons from the SGB. Firstly, we need to extract more nutrition and lifestyle data as possible from SGB. As of 2017, SGB also asked questions about tobacco use, additional eating, and physical activity behaviors. Momentarily, we are developing a shopping basket nutritional data base to increase the quantity of nutritional records as well as to mitigate the wishful thinking behavior, which is an important biased factor in the common surveys.

REFERENCES

[1] WHO, 2003. Diet, Nutrition, and the Prevention of Chronic Diseases. Report of a Joint WHO/FAO Ex-pert Consultation. *World Health Organization aper templates.*

[2] M. B. Schulze et al, Food based dietary patterns and chronic disease prevention, *BMJ 2018*; 361 doi: https://doi.org/10.1136/bmj.k2396, 13 June 2018

[3] Di Daniele, N.; Noce, A.; Vidiri, M.F.; Moriconi, E.; Marrone, G.; Annicchiarico-Petruzzelli, M.; D'Urso, G.; Tesauro, M.; Rovella, V.; De Lorenzo, A. Impact of Mediterranean diet on metabolic syndrome, cancer and longevity. *Oncotarget 2017*, 8, 8947–8979.

[4] De Lorenzo, A.; Noce, A.; Bigioni, M.; Calabrese, V.; Della Rocca, D.G.; Di Daniele, N.; Tozzo, C.; Di Renzo, L. The effects of Italian Mediterranean organic diet (IMOD) on health status. *Curr. Pharm. Des. 2010*, 16, 814–824.

[5] Andreoli, A.; Lauro, S.; Di Daniele, N.; Sorge, R.; Celi, M.; Volpe, S.L. Effect of a moderately hypoenergetic Mediterranean diet, and exercise program on body cell mass and cardiovascular risk factors in obese women. Eur. J. Clin. Nutr. 2008, 62, 892–897.

[6] Di Daniele, N.; Di Renzo, L.; Noce, A.; Iacopino, L.; Ferraro, P.M.; Rizzo, M.; Sarlo, F.; Domino, E.; De Lorenzo, A. Effects of Italian Mediterranean organic diet vs. low-protein diet in nephropathic patients according to MTHFR genotypes. *J. Nephrol. 2014*, 27, 529–536.

[7] Noce, A.; Marrone, G.; Urciuoli, S.; Di Daniele, F.; Di Lauro, M.; Pietroboni Zaitseva, A.; Di Daniele, N.; Romani, A. Usefulness of Extra Virgin Olive Oil Minor Polar Compounds in the Management of Chronic Kidney Disease Patients. *Nutrients 2021*, 13, 581.

[8] Noce, A.; Fabrini, R.; Bocedi, A.; Di Daniele, N. Erythrocyte glutathione transferase in uremic diabetic patients, *Acta Diabetol. 2015*, 52, 813–815.

[9] Di Marco, F.; Trevisani, F.; Vignolini, P.; Urciuoli, S.; Salonia, A.; Montorsi, F.; Romani, A.; Vago, R.; Bettiga, A. Preliminary Study on Pasta Samples Characterized in Antioxidant Compounds and Their Biological Activity on Kidney Cells. *Nutrients 2021*, 13, 1131.

[10] Koch, W. Dietary Polyphenols-Important Non-Nutrients in the Prevention of Chronic Noncommunicable Diseases. A Systematic Review. Nutrients 2019, 11, 39.

[11] Chen, Q.; Espey, M.G.; Krishna, M.C.; Mitchell, J.B.; Corpe, C.P.; Buettner, G.R.; Shacter, E.; Levine, M. Pharmacologic ascorbic acid concentrations selectively kill cancer cells: Action as a pro-drug to deliver hydrogen peroxide to tissues. *Proc. Natl. Acad. Sci.* USA 2005, 102, 13604–13609

[12] Noce, A.; Marrone, G.; Di Daniele, F.; Di Lauro, M.; Pietroboni Zaitseva, A.; Wilson Jones, G.; De Lorenzo, A.; Di Daniele, N. Potential Cardiovascular and Metabolic Beneficial Effects of omega-3 PUFA in Male Obesity Secondary Hypogonadism Syndrome. *Nutrients 2020*, 12, 2519.

[13] Owen, A.J.; Abramson, M.J.; Ikin, J.F.; McCaffrey, T.A.; Pomeroy, S.; Borg, B.M.; Gao, C.X.; Brown, D.; Liew, D. Recommended Intake of Key Food Groups and Cardiovascular Risk Factors in Australian Older, Rural-Dwelling Adults. *Nutrients 2020*, 12, 860.

[14] Kee, S. K., Son, Y. J, Kim H.G., Lee J. Il., Cho, H.S., Lee, S., 2014, Associations between food and bever-age groups and major diet-related chronic diseases: an exhaustive review of pooled/meta-analyses and systematic reviews, *Nutr Rev. 2014 Dec*; 72(12):741-62. doi: 10.1111/nure.12153

[15] Qudsi, D., Kartiwi, M., Saleh, N.B., 2017, Predictive data mining of chronic diseases using decision tree: A case study of health insurance company in Indonesia. *International Journal of Applied Engineer-ing Research 12(7)*:1334-1339

[16] Lei Z., Yang, S., Liu, H., Aslam, S., Liu, J., Bugingo, E., Zhang, D., 2018, Mining of Nutritional Ingredi-ents in Food for Disease Analysis, *IEEE Access 6(1)*:52766-52778

[17] McCabe, R.M, Adomavicius, G., Johnson P.E., Rund, E., Rush, A., Sperl-Hillen, A., 2008, Using Data Mining to Predict Errors in Chronic Disease Care, Advances in Patient Safety, *New Directions and Alternative Approaches in Vol. 3*: Performance and Tools.

[18] Haslam, D.W., James, W.P.T., Obesity, In the Lancet, Volume 366, Issue 9492, Pages 1197-1209

[19] Lange, K.W., James W.P.T., Makulska-Gertruda E., Nakamura Y., Reissmann, A., 2008, A. Sperl-Hillen, Using Data Mining to Predict Errors in Chronic Disease Care, Advances in Patient Safety. *New Directions and Alternative Approaches* (Vol. 3: Performance and Tools)

[20] Hearty, A.P., Gibney, M.J., 2008, A. Richonnet, C., Mazur, A., Analysis of meal patterns with the use of supervised data mining techniques—artificial neural networks and decision trees, *American Journal of Clinical Nutrition*, Volume 88, Issue 6, Pages 1632–1642.

[21] Von Ruesten, A., Feller, S., Bergmann, N.M., Boeing, H., 2013, S., Diet and risk of chronic diseases: results from the first 8 years of follow-up in the EPIC-Potsdam study, *European Journal of Clinical Nutrition volume 67*, pages412–419.

[22] Yu E. Y. W., Wesselius A., Sinhart C., Wolk A., 2020, A data mining approach to investigate food groups related to incidence of bladder cancer, *Bladder cancer Epidemiology and Nutritional Determinants International Study*, Cambridge University Press

[23] Einsele, F., Sadeghi, L., Ingold, R., Jenzer, H., 2015, A Study about Discovery of Critical Food Consumption Patterns Linked with Lifestyle Diseases using Data Mining Methods, *HealthInf, BIOSTEC - International Joint Conference on Biomedical Engineering Systems and Technologie*s, Lisbon.

[24] Mewes I., Jenzer H., Einsele, F., 2021, A Study about Discovery of Critical Food Consumption Patterns Linked with Lifestyle Diseases for Swiss Population using Data Mining Methods, *14th International Conference on Health Informatics*

[25] Lustenberger T., Jenzer H., Einsele F., 2022, Discovery of Association Rules of the Relationship between Food Consumption and Lifestyle Diseases from Swiss Nutrition's (MENUCH) Dataset & Multiple Swiss Health Datasets from 1992 To 2012, *6th International Conference on Big Data & Health BDHI*, AISCA, NET, DNLP, BDHI – 2022, pp. 77-93, 2022

AUTHORS

Farshideh Einsele, Prof. Dr., Lecturer & researcher in the Business section of BUAS, she teaches business informatic subjects and her research work is dedicated to epidemiology, big data and data mining

Jonas Baschung, Student in Bern University of Applied Sciences. He currently fulfilled his Bachelor in business information systems.

AN CONTEXT-AWARE INTELLIGENT SYSTEM TO AUTOMATE THE CONVERSION OF 2D AUDIO TO 3D AUDIO USING SIGNAL PROCESSING AND MACHINE LEARNING

Bolin Gao[1] and Yu Sun[2]

[1]Fairmont Preparatory Academy, 2200 W Sequoia Ave, Anaheim, CA 92801
[2]California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

ABSTRACT

As virtual reality technologies emerge, the ability to create immersive experiences visually drastically improved [1]. However, in order to accompany the visual immersion, audio must also become more immersive [2]. This is where 3D audio comes in. 3D audio allows for the simulation of sounds from specific directions, allowing a more realistic feeling [3]. At the present moment, there lacks sufficient tools for users to design immersive audio experiences that fully exploit the abilities of 3D audio.

This paper proposes and implements the following systems [4]:

1. Automatic separation of stems from the incoming audio file, or letting the user upload the stems themselves
2. A simulated environment in which the separated stems will be automatically placed in
3. A user interface in order to manipulate the simulated positions of the separated stems.
We applied our application to a few selected audio files in order to conduct a qualitative evaluation of our approach. The results show that our approach was able to successfully separate the stems and simulate a dimensional sound effect.

KEYWORDS

3D Audio, signal processing, Head Related Transfer Functions.

1. INTRODUCTION

3D audio is a powerful set of tools that allows users to virtually place sound effects in a three dimensional space by manipulating sounds coming from stereo or surround sound sources using head-related transfer functions and reverberations [5]. It is capable of tricking the brain into thinking that sounds are coming from different locations, while, for example, only using the two speakers present in headphones.

3D audio has been rising in popularity due to its ability to create immersive audio experiences, particularly for virtual reality applications, which are also gradually rising in popularity [6]. It is capable of simulating sounds from different locations without there being an actual audio source present. However, current usages are limited to sound effects within video games, and there does not seem to be enough tools specifically aimed at creating these immersive audio experiences.

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 123-133, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121610

Existing tools for manipulating 3D audio seem to be clunky and hard to use. For example, Dear Reality's dearVR exists as a plugin for digital audio workstations. While it is a powerful tool and contains many features, it is difficult to navigate, and the user interface is complicated and hard to use.

This paper seeks to address the ease-of-use issues of currently existing 3D audio manipulation technologies by automating the process of creating 3D audio experiences via pre-determined layouts, then granting the users a simple and intuitive interface for continued customization. By doing this, the process of creating immersive audio experiences using 3D audio will be greatly simplified, allowing more users to have access to its capacities and lower the learning curve.

Some of the existing techniques and systems for creating immersive 3D audio include tools like Dear Reality's DearVR, which allows users to control 3D audio in their sound production projects like music production or post production in filming. However, these tools assume the users to have professional experience with audio engineering, which is inapplicable to a large number of users. Others, such as Points2Sound, attempt to convert mono audio signals to 3D audio by using a deep learning model to interpret 3D visual information, then calculating the audio signals based on the predicted positions [7]. While their techniques are much easier to bring to users than Dear Reality's tools, their implementations are also limited in scale because of the requirement of existing 3D visual information, which might not always be available in cases where users only have auditory information.

In this paper, we follow the same line of research as others before us that created 3D audio, but with a focus on usability. Our goal is to facilitate the conversion of mono audio to 3D binaural audio using Head Related Transfer Functions in a convenient and effective fashion. Our method is inspired by previously existing tools that implement 3D audio, like Points2Sound and DearVR and their ability to allow users to create more immersive auditory experiences. These are the main features of our system: first, the ability to perform source separation on the input mono audio file into different stems. Second, the automatic conversion of the separated mono audio signals into 3D audio based on preset positional data. Third, a comprehensive and easy-to-use tool for manipulating the positional data of the separated input audio files. We believe that these parts would constitute an efficient system that could make the immersiveness offered by 3D audio more accessible, and potentially bring it to a wider audience.

In our application scenarios, we demonstrate how the above combination of techniques would work in practice. First, we show the usefulness of our approach by presenting some practice examples, which include different audio segments, and the subsequent results after being processed by our system. Afterwards, we then obtain samples of the processed, 3D audio then compare the different presets against each other to analyze the effectiveness of our system to create the desired 3D audio effects. By comparing the different 3D audio signals generated by the system, we should be able to determine the difference between them and use them to discern the actual direction for where the sound is coming from. If we are able to discern it, then our system is effective.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we met during the experiment and designing the sample; Section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 2; Section 4 presents the relevant details about the experiment we did, following by presenting the related work in Section 5. Finally, Section 6 gives the conclusion remarks, as well as pointing out the future work of this project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Requirement of individual audio stems

In order to create a spatialized sound effect, multiple stems of the audio would be needed. Else it would sound like all the inputted audio is coming from a singular direction. By having multiple audio stems, the program could simulate the effect of sounds coming from a multitude of directions. One option would be to ask for individual audio stem inputs in the function, however, they might not always be available, so we thought about a different approach: source separation. By using stem source separation technology, the singular audio source that users input can be separated into different stems automatically, allowing us to then apply the 3D audio sound effects to the different stems and place them into different virtual locations and create the desired immersion with headphones. For this project, we have opted to use Spleeter by Deezer, an open source source separation library written in python with state-of-the-art performance [8].

2.2. Application of 3D audio effects

With the separated audio stems, we then need a method to create the 3D sound effects on these stems. In order to solve this problem, we decided to utilize head-related transfer functions (HRTFs), which tells us how the left and right ears receive sounds from different locations differently [9]. When we apply these HRTFs to the audio stems and play them back through stereo headphones, sounds from specific locations and directions can be mimicked. By applying different HRTFs to different audio stems and merging them together, we can then create the effect of different sounds coming from different directions. We have decided to use the LISTEN HRTF database, measured at Ircam and AKG. It contains HRTFs measured from every direction surrounding the subject plus different elevations that would be useful for adding dimension to our 3D audio project. There are also a large number of subjects to choose from, which is more than enough for our purpose.

2.3. The customization of the output signal

After separating the audio stems, our system then applies 3D audio effects by using predetermined prefabs that put specific audio in specific locations. However, this does not exploit the full potential of 3D audio to create immersive sound designs. Our proposal to solve this problem is to add further functionality by allowing users to customize their input audio signal’s 3D position with an interactive UI element [15]. This will allow the users to dictate where they want the audio source to sound like they come from, allowing them to customize the 3D audio experience. Also, we will allow the users to add multiple input signals. This allows for the users to have more choice while uploading their audio file. For example, users can upload individual audio stems for better and more customizable positioning than what is offered by Spleeter, which only allows separation up to 5 predetermined stems, or they can merge completely different audio files together for an immersive experience.

3. SOLUTION

Our system allows for the creation of 3D audio from mono audio signals. First, the user will input a mono or stereo audio file. Then we separate the audio into different stems, for example, the vocal, piano, drums. We will also allow the user to input their own stems for further customization, as the automatic separated stems only allow up to 5 different stems. By letting the user upload their own stems allow more stems to be added for more flexibility. Afterwards, the system then calculates the new, specialized positions using the HRTFs and separated audio files. Now that the spatialized audio is created, we then implement a graphical user interface to allow for further customization of the 3D audio signal [13]. In order to achieve this goal, our system consists of the following components:

- A source separation algorithm for the separation of the input signal into respective stems
- A series of computations to automatically create the new specialized, 3D audio using HRTFs and the separated stems
- An interactive graphical interface to allow for further customization of the resulting 3D audio

The following diagram illustrates the components of our system.

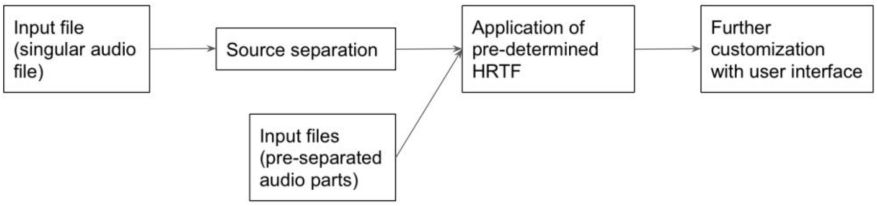

Input file (singular audio file)
Source separation
Application of pre-determined HRTF
Further customization with user interface
Input files (pre-separated audio parts)

Figure 1. Components of our system

In the next section, we will example and examine how each component works in depth.

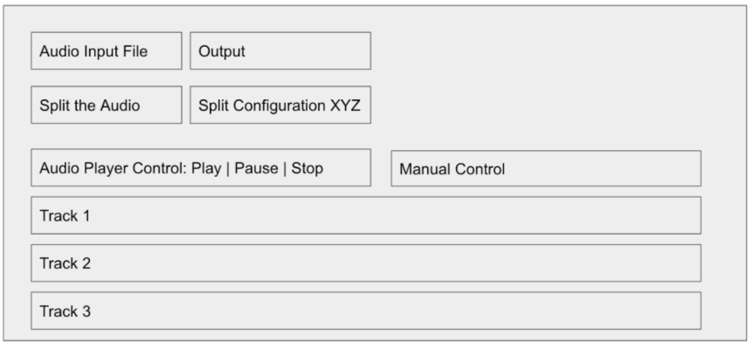


Figure 2. How components work

In this section we will explain how each component of the system works, starting with the source separation algorithm.

1. Source separation algorithm

The goal of the source separation algorithm is to splice the original, singular audio file into different parts so different HRTFs can be applied to them in order to create the effect that sounds are coming from different locations, therefore creating the 3D audio effect. In order to accomplish the source separation, we will use a pre-trained algorithm called Spleeter, made by Deezer. This algorithm is capable of 2-stem, 4-stem, and 5-stem source separation. It uses the original input file and uses pre-trained algorithms to extract parts such as the vocals, drums, and piano. With this feature, we can now apply directional information to the separated stems.

2. Application of directional information to the audio stems

Now that we have the audio stems, we need a way to create 3D audio effects by applying directional information to each of the stems to make them sound like they are coming from different directions, simulating the effects of being present in a 3D space and hearing sounds from different locations. To accomplish this effect, we will be applying HRTFs to the different audio stems. HRTFs are pre-recorded from different directions that show how the two ears react to them differently. By applying these HRTFs to our audio stems, we can calculate how the audio stems would sound like if they are from these directions. For our system, we will be using the LISTEN HRTF database, which contains HRTFs for all 360 degrees from the listener, along with different elevations above and below the listener for a larger number of options. We will automatically apply the HRTFs of specific directions to the audio stems to automatically create the 3D audio effects, then allow the user to further customize the resulting 3D audio signal with an interactive user interface.

3. Graphical user interface to allow for further signal customization

While we already successfully generated a 3D signal automatically, we have yet to access the full potential of 3D audio. In order to do that, we will allow the user to further customize their 3D audio using a graphical user interface. In the graphical user interface, the user can determine the simulated location they want each of their audio stems to come from. We will also add the ability for users to upload their own audio stems and directly apply the 3D audio to them. This allows users to directly apply positional data to their own stem, allowing for possibilities beyond the 2 stems, 4 stems, 5 stem separations offered by Spleeter.

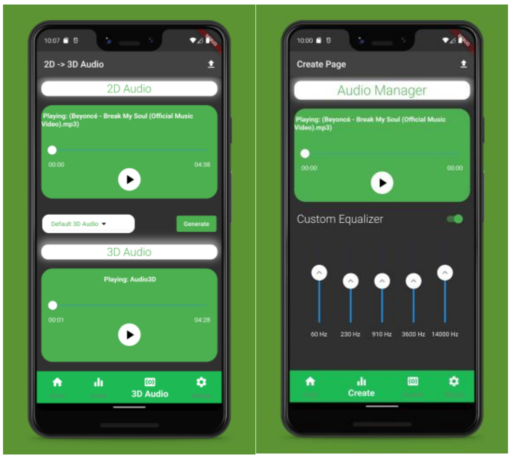


Figure 3. GUI for customizing 3D audio

4. EXPERIMENT

4.1. Experiment 1

To test the effectiveness of our system, we have decided to test it on two different real use case situations in the following procedure:

1. Grab the audio file
2. Use mat lab to plot the functions
3. Apply the HRTF to it
4. Plot it again and compare their differences
5. Add the audio files too for listening purposes

One case where we use Spleeter and one case where we just upload the parts.

Test a few different angles of the thing.

In order to test the effectiveness of our system, we decided to conduct an experiment to see how the system functions under a practical situation.

Here is the original audio file we are testing, it is composed of a drum track, a trumpet track, pianos, and a variety of other accompaniment in the background:

Embed the wav file here

Here is the Time Domain visualization of the input signal

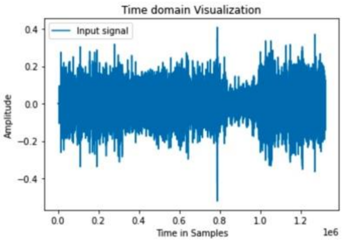


Figure 4. Time Domain visualization

And here is the Frequency Domain visualization of the input signal

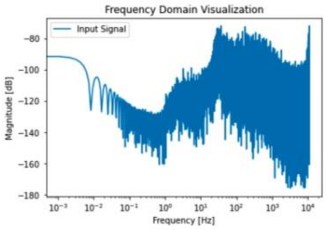


Figure 5. Frequency Domain visualization

For the first step of our system, we use the source separation algorithm in order to split the singular input file into multiple stems. We will utilize the 5 stem option in Spleeter from split the audio into 5 distinctive stems. Spleeter is capable of separating the track into the vocals, piano, drums, and bass. It will put the rest of the audio that does not fall into the previous sections into one signal called “other”.

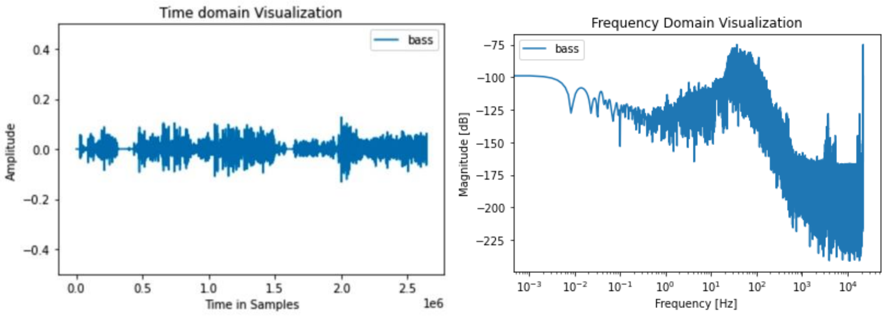


Figure 6. Time and frequency domain for bass

For the bass track, we can see that the amplitude is lower than the amplitude of the original audio signal. This means that the source separation algorithm was able to successfully separate parts of the original out of it. We can also see that this section does not contain much frequency of the higher ranges, notably from 100 to 1000 Hz. At the same time, the magnitude of the lower frequencies seem to be high, indicating that it successfully captured the bass of the original mix. There seems to be a peak at the very high frequencies, though. This is assumed to be an artifact from the source separation algorithm.

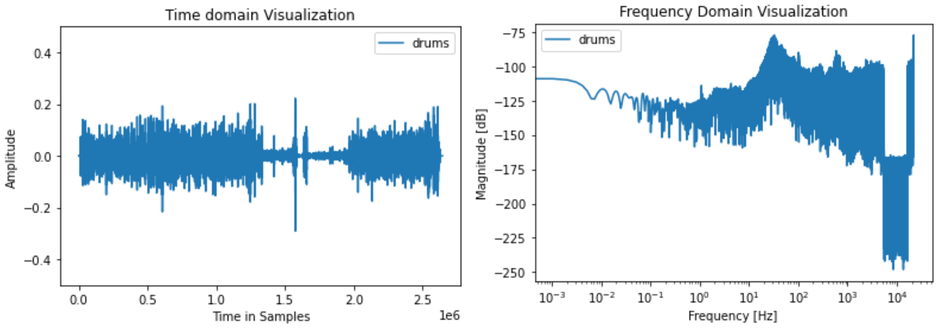


Figure 7. Time and frequency domain analysis for drums

The drum occupied a large amount of the original mix, so the amplitude seems to be higher than the amplitude on the bass most of the time. The frequency domain contains a large gap around 1000 Hz. We do not know the cause of this, and it seems like another artifact of the source separation algorithm. The drums contain a relatively large amount of magnitude at the very high frequencies after the gap, though. This is interpreted as the very high frequencies present in the original audio signal, as the high-hats are capable of generating these frequencies, unlike anything in the bass.

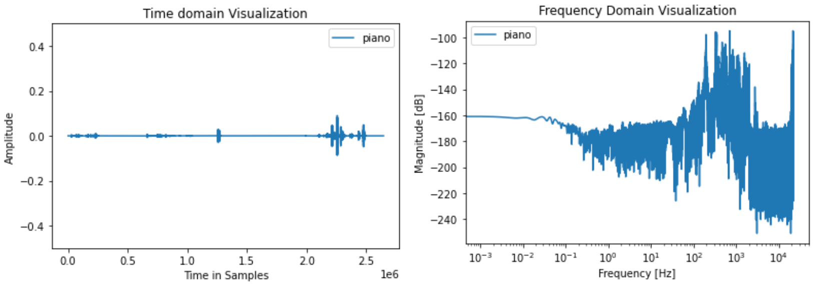


Figure 8. Time and frequency domain analysis for piano

The piano stem does not contain much of the original mix. The stem separation algorithm was not able to separate much of the piano out of the mix.

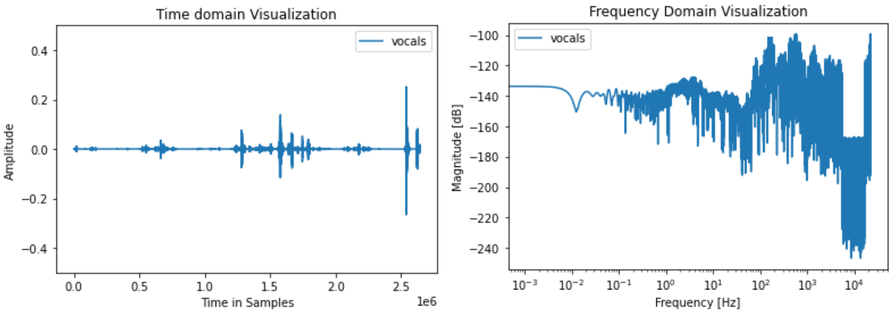


Figure 9. Time and frequency domain analysis for vocal

There should not be anything in the vocal stem, as the input audio does not have vocals. This is evident in the low amplitudes throughout the song. The signals that do exist here are artifacts of the source separation algorithm.

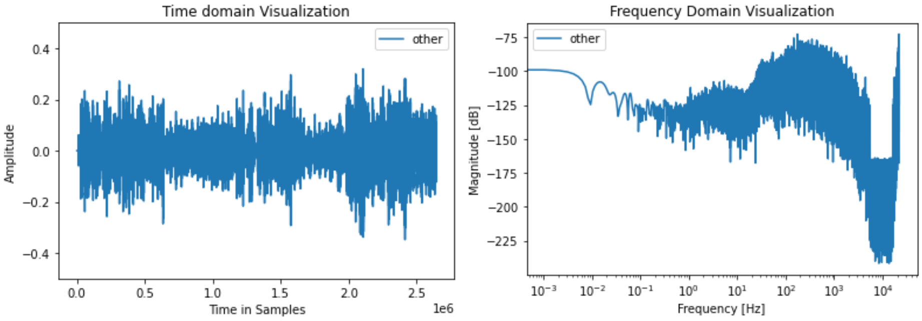


Figure 10. Time and frequency domain analysis for other instruments

Most of the instruments of the original audio falls outside of the category, so this stem contains much of the original audio. This is evident, as the amplitude of this stem is the highest. This distinctive gap at around 1000 Hz is still present.

Next, we will apply the HRTFs to the separated audio stems. In this example we will select four random HRTFs to the audio stems.

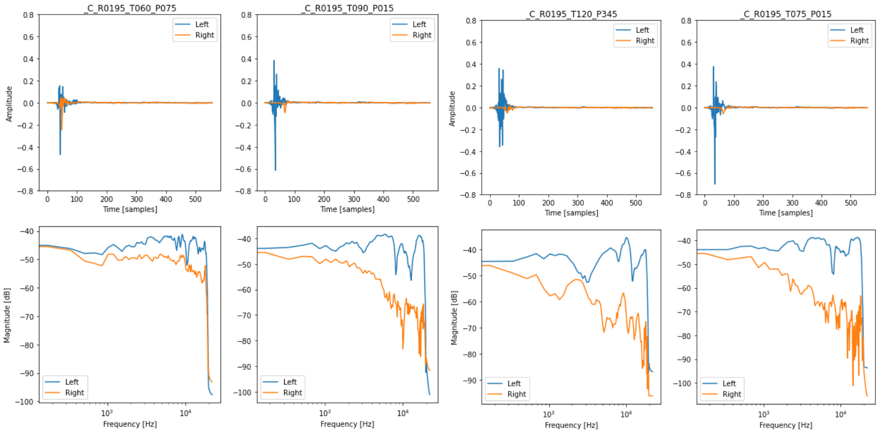


Figure 11. Time and frequency domain of the four HRTFs that are selected at random

We can see that all of these four amplitudes happen to be on the left, so we would expect more audio to be on the left side of the final mix.

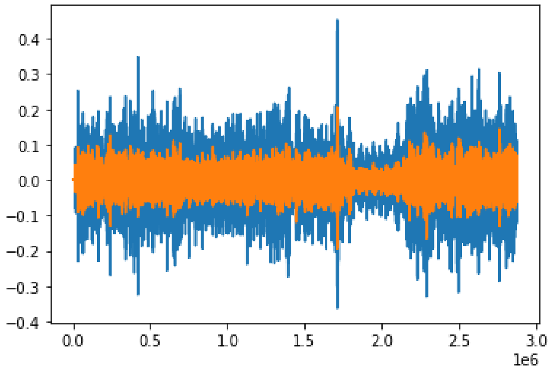


Figure 12. Amplitude of the left and right channels for the final mix

And that is confirmed by this graph showing the final amplitude.

6. Related Work

In this paper, the authors demonstrated a model that generates a binaural version of 2D audio based on 3D visual information from a scene [7]. This is comparable to our work, as we are both attempting to convert between 2D and 3D audio. However, their approach relies on visual information to determine how the generated binaural audio sounds. In our approach, we decided that rather than relying on visual cues, allowing direct customization of the output 3D signal with an interactive interface would allow for more freedom and generate more satisfactory results.

This paper presented a synthesis process in order to generate binaural audio from a mono audio source, also similar to what we are trying to accomplish [10]. Instead of utilizing head-related transfer functions, they created a novel process for generating binaural audio utilizing diffusion models. Their system works very well compared to the HRTFs that we have used. As their system does not contain a system for further customization of the resulting signal. It would be nice if the two approaches can be connected together.

This paper showed a method for generating binaural audio with a deep neural network instead of HRTFs [11]. They also included a system to extract positional information from images, and with the model, they were also able to estimate a depth map to aid the generation of the final binaural signal, which is something we have not implemented in our system. We would like to implement the depth aspect of their model in the future. With this, the interface for customizing the final binaural signal could be further improved upon.

7. Conclusions

In this paper, we have proposed a system that enables the conversion of 2D to 3D audio. The system consists of the following components: the separation of the input audio file, the application of HRTFs to the audio file, and an interface for further customization of the output 3D audio file [12]. For the separation of the input file, we have utilized Splitter, a source separation algorithm that's capable of 2, 4, and 5 stem separation. After separating the input audio into stems, we then apply different HRTFs to each of the stems. For this, we will utilize the Listen HRTF database. After applying the predetermined configurations to the audio stems, we then allow the users to further customize their output file with a graphical interface that allows users to apply different HRTFs to each stem.

To test the effectiveness of our system, we have applied it to different audio files to determine its performance. First, we analyzed whether or not the source separation algorithm was able to successfully accomplish its goal. Then we applied different HRTFs to the audio sources and analyzed its effectiveness. We then analyzed the time and frequency domains of the files at different stages. Based on the results of our experiments, we were able to see that the source separation algorithm and the application of HRTF has worked. And with this, we have concluded that our system is effective and was able to accomplish our goals.

For our current method, there exists a range of limitations that we have not yet been able to address. First, the source separation algorithm is only capable of performing source separation for a few pre-determined types of stems, such as the vocals, drums, and bass. Within these confinements, it is able to perform quite well, however, more options for separation will allow for more possibilities of customization. Another limitation of the current method is that it is unable to simulate the distance from the virtual audio source. If the distance is able to be simulated, then the virtual environment of the 3D audio would be complete. We would also like to add a way to dynamically apply HRTFs to the audio stems to make them appear to be moving through a 3D environment [14].

For the source separation algorithm, we hope to improve on it by further training the algorithm to obtain better results, and extend the capabilities of it to allow for further separation of different components of the input audio file. For the distance from the virtual audio source problem, we can manipulate the audio to approximate the distance. We would also like to have a method to make the audio files to appear as if they are moving through space.

REFERENCES

[1] Boas, Y. A. G. V. "Overview of virtual reality technologies." Interactive Multimedia Conference. Vol. 2013. 2013.

[2] Cummings, James J., and Jeremy N. Bailenson. "How immersive is enough? A meta-analysis of the effect of immersive technology on user presence." Media psychology 19.2 (2016): 272-309.

[3] Brinkman, Willem-Paul, Allart RD Hoekstra, and René van EGMOND. "The effect of 3D audio and other audio techniques on virtual reality experience." Annual Review of Cybertherapy and Telemedicine 2015 (2015): 44-48.

[4] Sundareswaran, Venkataraman, et al. "3D audio augmented reality: implementation and experiments." The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.. IEEE, 2003.

[5] Frauenberger, Christopher, and Markus Noistering. "3D audio interfaces for the blind." Georgia Institute of Technology, 2003.

[6] Sherman, William R., and Alan B. Craig. "Understanding virtual reality." San Francisco, CA: Morgan Kauffman (2003).

[7] Lluís, Francesc, Vasileios Chatziioannou, and Alex Hofmann. "Points2Sound: From mono to binaural audio using 3D point cloud scenes." arXiv preprint arXiv:2104.12462 (2021).

[8] Hennequin, Romain, et al. "Spleeter: a fast and efficient music source separation tool with pre-trained models." Journal of Open Source Software 5.50 (2020): 2154.

[9] Brown, C. Phillip, and Richard O. Duda. "An efficient HRTF model for 3-D sound." Proceedings of 1997 Workshop on Applications of Signal Processing to Audio and Acoustics. IEEE, 1997.

[10] Leng, Yichong, et al. "BinauralGrad: A Two-Stage Conditional Diffusion Probabilistic Model for Binaural Audio Synthesis." arXiv preprint arXiv:2205.14807 (2022).

[11] Parida, Kranti Kumar, Siddharth Srivastava, and Gaurav Sharma. "Beyond mono to binaural: Generating binaural audio from mono audio with depth and cross modal attention." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2022.

[12] Hornstein, Jonas, et al. "Sound localization for humanoid robots-building audio-motor maps based on the HRTF." 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2006.

[13] Guizzo, Eric, et al. "L3DAS21 Challenge: Machine learning for 3D audio signal processing." 2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2021.

[14] Azim, Asma, and Olivier Aycard. "Detection, classification and tracking of moving objects in a 3D environment." 2012 IEEE Intelligent Vehicles Symposium. IEEE, 2012.

[15] Sermuga Pandian, Vinoth Pandian, Sarah Suleri, and Prof Dr Matthias Jarke. "UISketch: a large-scale dataset of UI element sketches." Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021.

QUANTIFYING THE THEORY VS. PROGRAMMING DISPARITY USING SPECTRAL ANALYSIS

Natarajan Meghanathan

Department of Electrical & Computer Engineering and Computer Science,
Jackson State University, Jackson, MS, USA

ABSTRACT

Some students in the Computer Science and related majors excel very well in programming-related assignments, but not equally well in the theoretical assignments (that are not programming-based) and vice-versa. We refer to this as the "Theory vs. Programming Disparity (TPD)". In this paper, we propose a spectral analysis-based approach to quantify the TPD metric for any student in a course based on the percentage scores (considered as decimal values in the range of 0 to 1) of the student in the course assignments (that involves both theoretical and programming-based assignments). For the student whose TPD metric is to be determined: we compute a Difference Matrix of the scores in the assignments, wherein an entry (u, v) in the matrix is the absolute difference in the decimal percentage scores of the student in assignments u and v. We subject the Difference Matrix to spectral analysis and observe that the assignments could be partitioned to two disjoint sets wherein the assignments within each set have the decimal percentage scores closer to each other, and the assignments across the two sets have the decimal percentage scores relatively more different from each other. The TPD metric is computed based on the Euclidean distance between the tuples representing the actual numbers of theoretical and programming assignments vis-a-vis the number of theoretical and programming assignments in each of the two disjoint sets. The larger the TPD score (in a scale of 0 to 1), the greater the disparity and vice-versa.

KEYWORDS

Spectral Analysis, Theory vs. Programming Disparity, Eigenvector, Bipartivity.

1. INTRODUCTION

Spectral analysis of a complex network has been observed to reveal significant structural details that would have been hitherto unknown in the scientific community [1]. Spectral analysis of a network graph typically involves the computation of the Eigenvectors and their corresponding Eigenvalues using one of the symmetric matrices such as the adjacency matrix, Laplacian matrix [7] and etc that reflect the topology of the network [2]. The indexes of the entries in an Eigenvector correspond to the sorted order (increasing order) of the node ids. Spectral analysis of an *nxn* matrix results in *'n'* Eigenvalues and the corresponding 'n' Eigenvectors (one Eigenvector per Eigenvalue). Any two Eigenvectors are orthogonal to each other (i.e., the dot product of any two Eigenvectors is zero). The largest Eigenvalue is called the "Principal Eigenvalue" and the Eigenvector corresponding to the principal Eigenvalue is called the "Principal Eigenvector" [3]. If all the entries in the underlying matrix used for spectral analysis are positive (≥ 0), then the principal Eigenvalue is guaranteed to be positive and all the entries in the principal Eigenvector are also positive. Hence, when computed based on matrices with positive entries, in order to still

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 135-145, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121611

satisfy the *mutually orthogonal property*, (unless all the entries in an Eigenvector are zero) at least one non-zero entry in every Eigenvector other than the Principal Eigenvector is guaranteed to be negative so that the dot product of any two Eigenvectors evaluates to zero.

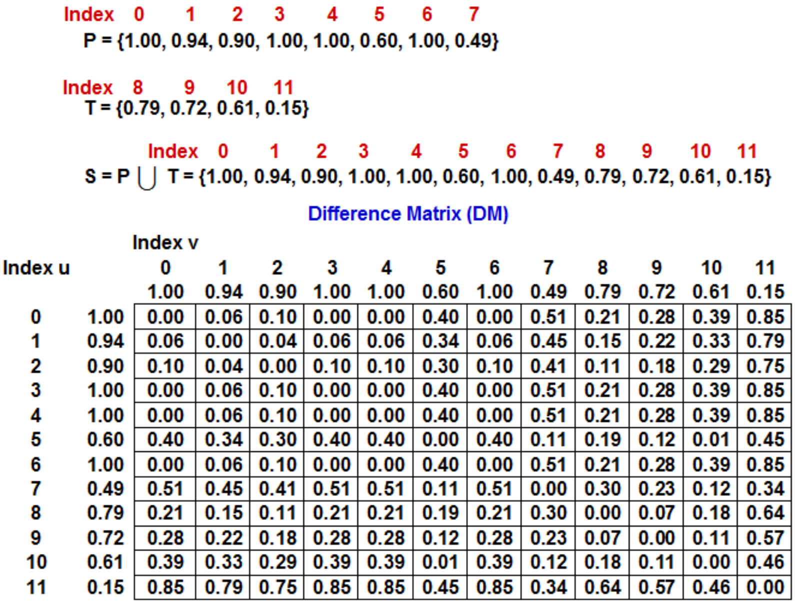
In a classical work [4], Estrada et al proposed that the extent of bipartivity (in the form of a bipartivity index) among the vertices in a network could be quantified using the Eigenvalues of the adjacency matrix of the network graph. A graph is said to be bipartite if the vertices of the graph could be grouped into two disjoint sets such that the end vertices of any edge are in the two different partitions and not in the same partition. Estrada et al observed that if the underlying graph is bipartite, the vertices with positive and negative entries in the Eigenvector (hereafter referred to as the 'bipartite Eigenvector') corresponding to the smallest Eigenvalue represent the two disjoint partitions of vertices in the graph. Estrada et al also observed that if the underlying graph is not bipartite, the vertices with positive and negative entries in the bipartite Eigenvector could still be construed to form the two disjoint partitions of the vertices of the graph such that there are exist a minimal number of edges (referred to as the 'frustrated edges') between vertices in the same partition and a majority of the edges are between vertices across the two partitions.

In this paper, we conduct spectral bipartivity analysis of the scores earned by a student in theoretical and programming assignments of a Computer Science course and seek to quantify the extent of disparity in the scores earned by the student in the theoretical assignments vs. programming assignments. Some students in the Computer Science and related majors excel very well in programming-related assignments, but not equally well in the theoretical assignments (that are not programming-based) and vice-versa. We refer to this as the "Theory vs. Programming Disparity (*TPD*)". Our methodology is briefly described here (more details are in Section 2): The student score in each assignment is considered in a decimal percentage scale of 0 to 1 (i.e., each assignment is evaluated for 100% and the decimal percentage score for a student in the assignment is the percentage score divided by 100: for example, if an assignment score is 81%, the decimal percentage score is 81/100 = 0.81). We first determine the Difference Matrix (*DM*) of the student scores in the assignments wherein an entry DM_{uv} is the absolute difference in the decimal percentage scores of the two assignments *u* and *v*. We then determine the bipartite Eigenvector of the *DM* by subjecting it to spectral analysis. We identify the index entries with positive signs and negative signs, and the corresponding assignment IDs are grouped into two separate (disjoint) sets. We observe that any two assignments with similar (closer) values for the decimal percentage scores are more likely to be grouped into the same set and two assignments with appreciably different decimal percentage scores are more likely to be grouped in separate sets. That is any two vertices *u* and *v* whose DM_{uv} entry is closer to 0 are more likely to end up in the same set of vertices and vertices *u* and *v* whose DM_{uv} entry is much greater than 0 are more likely to end up in different sets of vertices. Such an observation is consistent with the observations made by Estrada et al for bipartivity analysis using an adjacency matrix *A* (i.e., vertices *u* and *v* whose A_{uv} entries were 0 are more likely to be in the same partition and vertices *u* and *v* whose A_{uv} entries were 1 are more likely in different partitions). We quantify the *TPD* on the basis of the Euclidean distance between the actual number of theoretical and programming assignments vs. the number of theoretical and programming assignments in the two sets of disjoint assignments identified through spectral bipartivity analysis.

The rest of the paper is organized as follows: In Section 2, we present our proposed methodology to quantify the *TPD* metric using a running example. Section 3 evaluates the effectiveness of the proposed *TPD* approach with two of the well-known metrics (Bipartivity index: *BPI* [4] and Hausdorff Distance: *HD* [5]) that exist in the literature to study the effectiveness of partitioning of a data set to two clusters. Section 3 also highlights the uniqueness of the proposed *TPD* approach. Section 4 reviews related work and Section 5 concludes the paper. Throughout the paper, the terms 'set' and 'partition', 'network' and 'graph', 'edge' and 'link' are used interchangeably. They mean the same.

2. Spectral Bipartivity Analysis to Quantify Theory vs. Programming Disparity

Let *P* and *T* be respectively the set of scores (represented in decimal percentage format) earned by a student in programming and theoretical assignments. The indexes for the assignments in the set *P* range from 0 to |*P*| - 1, where |*P*| is the cardinality of the set *P* (i.e., the number of programming assignments). The indexes for the assignments in the set *T* range from |*P*| to |*P*| + |*T*| - 1, where |*T*| is the cardinality of the set *T* (i.e., the number of theoretical assignments). Let S be the union of the two sets *P* and *T*. That is, the set S comprises of the scores earned by the student in the programming assignments, followed by the theoretical assignments. The indexes for the assignments in *S* are the same as their indexes in the sets *P* or *T*, whichever they come from. Let *DM* be a symmetric/square matrix whose dimensions correspond to the cardinality of the set *S*. An entry DM_{ij} for row index *i* and column index *j* represents the absolute difference in the decimal percentage scores for the i^{th} and j^{th} element/assignment in the set *S*. Figure 1 presents the sets *P*, *T* and *S* (of size 8, 4 and 12 respectively) as well as the *DM* matrix (of dimensions 12 x 12) for a sample data set that is used as a running example to explain the proposed methodology in this section. The indexes for the assignments in the sets *P* and *T* range from 0...7 and 8...11 respectively; the indexes of these assignments are retained in the set *S* that is the amalgamation of the assignments in the sets *P* and *T* (in the same order). Likewise, the indexes in *DM* correspond to the indexes in the set *S*.



Index 0 1 2 3 4 5 6 7
P = {1.00, 0.94, 0.90, 1.00, 1.00, 0.60, 1.00, 0.49}

Index 8 9 10 11
T = {0.79, 0.72, 0.61, 0.15}

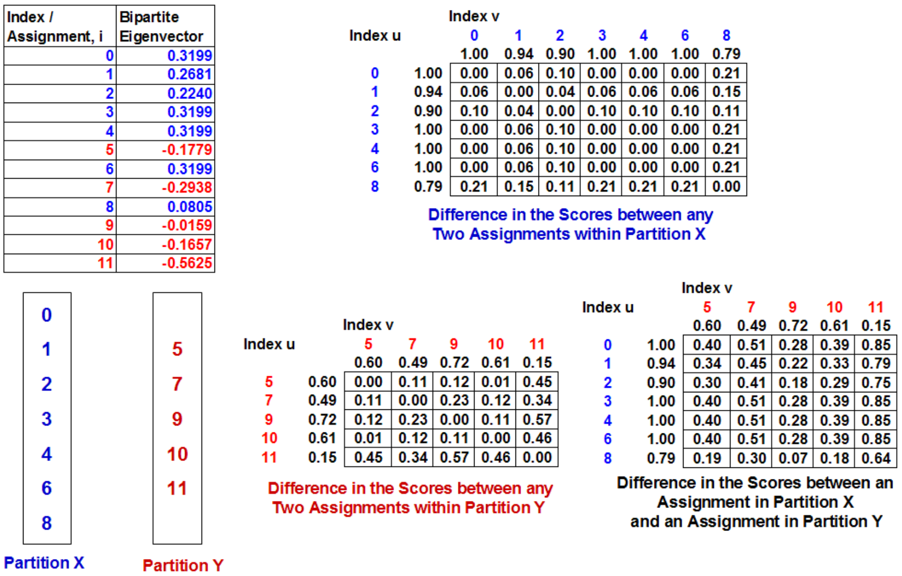
Index 0 1 2 3 4 5 6 7 8 9 10 11
S = P ⋃ T = {1.00, 0.94, 0.90, 1.00, 1.00, 0.60, 1.00, 0.49, 0.79, 0.72, 0.61, 0.15}

Difference Matrix (DM)

		Index v											
Index u		0	1	2	3	4	5	6	7	8	9	10	11
		1.00	0.94	0.90	1.00	1.00	0.60	1.00	0.49	0.79	0.72	0.61	0.15
0	1.00	0.00	0.06	0.10	0.00	0.00	0.40	0.00	0.51	0.21	0.28	0.39	0.85
1	0.94	0.06	0.00	0.04	0.06	0.06	0.34	0.06	0.45	0.15	0.22	0.33	0.79
2	0.90	0.10	0.04	0.00	0.10	0.10	0.30	0.10	0.41	0.11	0.18	0.29	0.75
3	1.00	0.00	0.06	0.10	0.00	0.00	0.40	0.00	0.51	0.21	0.28	0.39	0.85
4	1.00	0.00	0.06	0.10	0.00	0.00	0.40	0.00	0.51	0.21	0.28	0.39	0.85
5	0.60	0.40	0.34	0.30	0.40	0.40	0.00	0.40	0.11	0.19	0.12	0.01	0.45
6	1.00	0.00	0.06	0.10	0.00	0.00	0.40	0.00	0.51	0.21	0.28	0.39	0.85
7	0.49	0.51	0.45	0.41	0.51	0.51	0.11	0.51	0.00	0.30	0.23	0.12	0.34
8	0.79	0.21	0.15	0.11	0.21	0.21	0.19	0.21	0.30	0.00	0.07	0.18	0.64
9	0.72	0.28	0.22	0.18	0.28	0.28	0.12	0.28	0.23	0.07	0.00	0.11	0.57
10	0.61	0.39	0.33	0.29	0.39	0.39	0.01	0.39	0.12	0.18	0.11	0.00	0.46
11	0.15	0.85	0.79	0.75	0.85	0.85	0.45	0.85	0.34	0.64	0.57	0.46	0.00

Figure 1. Sample Data set to Illustrate the Spectral Bipartivity-based Analysis for Theory vs. Programming Disparity

Figure 2 presents the 12 Eigenvalues and the entries of the Bipartite Eigenvector, which is the Eigenvector corresponding to the smallest Eigenvalue of -2.2285, of the Difference Matrix (*DM*). The indexes (colored in blue) whose entries are positive (≥ 0) are grouped to partition (set) *X* and the indexes (colored in red) whose entries are negative (< 0) are grouped to partition (set) *Y*. These are the two disjoint partitions of the assignments in the set *S* predicted based on spectral bipartivity analysis. In Figure 2, we also display the submatrices of the Difference Matrix that show the difference in the decimal percentage scores for any two assignments within the sets *X* and *Y* as well as between the two sets *X* and *Y*. We notice the entries in the submatrices corresponding to the differences in the assignment scores within the sets *X* or *Y* are relatively much smaller (closer to 0) compared to the entries in the submatrix corresponding to the differences in the assignment scores between an assignment in set *X* and an assignment in set *Y*.



Index / Assignment, i	Bipartite Eigenvector
0	0.3199
1	0.2681
2	0.2240
3	0.3199
4	0.3199
5	-0.1779
6	0.3199
7	-0.2938
8	0.0805
9	-0.0159
10	-0.1657
11	-0.5625

Index u		Index v 0	1	2	3	4	6	8
		1.00	0.94	0.90	1.00	1.00	1.00	0.79
0	1.00	0.00	0.06	0.10	0.00	0.00	0.00	0.21
1	0.94	0.06	0.00	0.04	0.06	0.06	0.06	0.15
2	0.90	0.10	0.04	0.00	0.10	0.10	0.10	0.11
3	1.00	0.00	0.06	0.10	0.00	0.00	0.00	0.21
4	1.00	0.00	0.06	0.10	0.00	0.00	0.00	0.21
6	1.00	0.00	0.06	0.10	0.00	0.00	0.00	0.21
8	0.79	0.21	0.15	0.11	0.21	0.21	0.21	0.00

Difference in the Scores between any Two Assignments within Partition X

Partition X	Partition Y
0	
1	5
2	7
3	9
4	10
6	11
8	

Index u		Index v 5	7	9	10	11
		0.60	0.49	0.72	0.61	0.15
5	0.60	0.00	0.11	0.12	0.01	0.45
7	0.49	0.11	0.00	0.23	0.12	0.34
9	0.72	0.12	0.23	0.00	0.11	0.57
10	0.61	0.01	0.12	0.11	0.00	0.46
11	0.15	0.45	0.34	0.57	0.46	0.00

Difference in the Scores between any Two Assignments within Partition Y

Index u		Index v 5	7	9	10	11
		0.60	0.49	0.72	0.61	0.15
0	1.00	0.40	0.51	0.28	0.39	0.85
1	0.94	0.34	0.45	0.22	0.33	0.79
2	0.90	0.30	0.41	0.18	0.29	0.75
3	1.00	0.40	0.51	0.28	0.39	0.85
4	1.00	0.40	0.51	0.28	0.39	0.85
6	1.00	0.40	0.51	0.28	0.39	0.85
8	0.79	0.19	0.30	0.07	0.18	0.64

Difference in the Scores between an Assignment in Partition X and an Assignment in Partition Y

Figure 2. Bipartite Eigenvector for the Difference Matrix of Figure 1 as well as the Submatrices Representing the Difference Values between Assignments within the two Partitions and across the Two Partitions

We now seek to quantify the extent to which the grouping of the assignments in sets *X* and *Y* are closer to the grouping of the programming and theoretical assignments in the sets *T* and *P*. In order for our hypothesis (that there exists a disparity in the scores earned by the students in the programming vs. theoretical assignments) to be true, we would like the two sets *X* and *Y* to be the same as the two sets *T* and *P* (or *P* and *T*); that is, we would prefer the set *X* to be all theoretical assignments and the set *Y* to be all programming assignments or the set *X* to be all programming assignments and the set *Y* to be all theoretical assignments. In this pursuit, we first determine the number of theoretical assignments and the number of programming assignments in each of the sets *X* and *Y* and let these be indicated using symbols $|X_T|$, $|X_P|$, $|Y_T|$ and $|Y_P|$. We then determine the Euclidean distance between the tuples ($|X_T|$, $|Y_P|$) and ($|T|$, $|P|$) as well as between the tuples ($|Y_T|$, $|X_P|$) and ($|T|$, $|P|$) and refer to the minimum of these two Euclidean distances as the Theory vs. Programming Tuple Proximity (*TPTP*) distance for the given data set.

The maximum value for the *TPTP* distance is $\sqrt{|T|^2 + |P|^2}$ and we will incur it when all the four values $|X_T|$, $|X_P|$, $|Y_T|$ and $|Y_P|$ are zero each (i.e., the assignment IDs in the partitions *X* and *Y* identified through spectral bipartivity analysis have no overlap with the assignment IDs in the partitions *P* and *T*). Such a scenario occurs when there is minimal or no disparity among the scores in the programming vs. theoretical assignments and the distribution of the assignment IDs in the partitions *X* and *Y* are random. On the other hand, if there is maximum disparity, the assignment IDs in partitions *X* and *Y* will overlap with those of *P* and *T*, and either $|X_P| \approx |P|$ and $|Y_T| \approx |T|$ or $|Y_P| \approx |P|$ and $|X_T| \approx |T|$. This would make the *TPTP* distance much smaller than the maximum value of $\sqrt{|T|^2 + |P|^2}$. Considering the above interpretation of the *TPTP* distance, we formulate the *TPD* (Theory vs. Programming Disparity) metric as follows:

$$TPD = 1 - \frac{TPTP}{\sqrt{|T|^2 + |P|^2}}$$

If there is maximum disparity, then the *TPTP* distance will be either closer to 0 or the ratio *TPTP* / $\sqrt{|T|^2 + |P|^2}$ be closer to 0, making the *TPD* metric score to be closer to 1. On the other hand, if there is no disparity, the *TPTP* distance will be closer to the maximum value of $\sqrt{|T|^2 + |P|^2}$ and as a result the ratio *TPTP* / $\sqrt{|T|^2 + |P|^2}$ be closer to 1, making the *TPD* metric score to be closer to 0.

Figure 3 presents the computation of the $|X_T|$, $|X_P|$, $|Y_T|$ and $|Y_P|$ numbers for the running example of Figures 1-2 and the computation of the *TPD* metric for the sample data set. The *TPD* metric value for this data set is 0.75, indicating that there is an appreciable disparity in the theoretical vs. programming scores in this data set. Such a conclusion could also be justified by visually looking at the proximity of the tuple ($|Y_T|$ = 3, $|X_P|$ = 6) corresponding to the *TPTP* distance to the tuple ($|T|$ = 4, $|P|$ = 8) in Figure 3 as well as the raw data set values {0.79, 0.72, 0.61, 0.15} and {1.00, 0.94, 0.90, 1.00, 1.00, 0.60, 1.00, 0.49} for the sets *T* and *P* respectively.

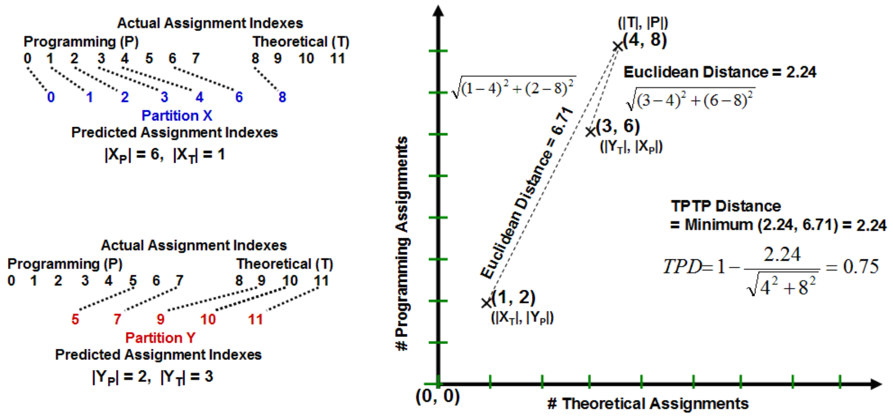


Figure 3. Computation of the *TPD* Metric for the Sample Data Set of Figure 1

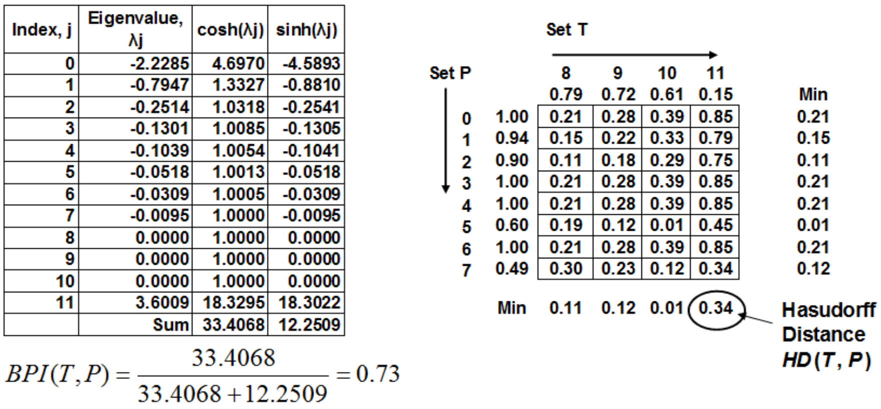
3. Evaluation of the Proposed Approach

In this section, we apply the proposed spectral bipartivity analysis-based approach to quantify the theory vs. programming disparity per student in the CSC 228 Data Structures and Algorithms course taught in Spring 2020 at Jackson State University, MS, USA. In addition to the *TPD* metric, we consider two other metrics that appear to be potentially applicable to quantify the extent of disparity in a data set with respect to two different categories (in this case, theoretical vs. programming assignments). These are:

(i) Bipartivity Index (*BPI*): The *BPI* was originally proposed by Estrada et al [4] to quantify the extent of bipartivity between the two partitions of vertices identified using the Eigenvector (referred to as the Bipartite Eigenvector) corresponding to the smallest Eigenvalue. The input matrix for Estrada et al's spectral bipartivity analysis is a 0-1 adjacency matrix. If the underlying graph is not bipartite, the two partitions of vertices identified using the Bipartite Eigenvector have as few edges as possible between vertices within the same partition and a majority of the edges are between vertices across the two partitions. If the underlying graph is indeed bipartite, the two partitions of vertices identified using the Bipartite Eigenvector will have no edges between vertices within the same partition and all the edges in the graph will be between vertices across the two partitions. The *BPI* of a graph of *'n'* vertices is computed using the following formulation based on the 'n' Eigenvalues (λ) of its 0-1 adjacency matrix. If the underlying graph is bipartite, the sum of the sinh values of the '*n*' Eigenvalues will be zero and the *BPI* of the graph will be 1.0. If the underlying graph is not bipartite, then the sum of the sinh values of the '*n*' Eigenvalues will be greater than 0, and the *BPI* will be less than 1.0.

$$BPI = \frac{\sum_{i=0}^{n-1} \cosh(\lambda_i)}{\sum_{i=0}^{n-1} \cosh(\lambda_i) + \sum_{i=0}^{n-1} \sinh(\lambda_i)}$$

In this section, we will explore how well the Eigenvalues of the Difference Matrix (*DM*) of an assignment score data set capture the Theory vs. Programming Disparity such that the *BPI* values are closer to 1.0 for data sets with larger values for the *TPD* metric and vice-versa. Figure 4 displays the 12 Eigenvalues of the running example assignment scores data set of Section 2: the sums of the cosh and sinh functions of the Eigenvalues are 33.4068 and 12.2509 respectively, leading to a *BPI* of 33.4068 / (33.4068 + 12.2509) = 0.73.



Index, j	Eigenvalue, λj	cosh(λj)	sinh(λj)
0	-2.2285	4.6970	-4.5893
1	-0.7947	1.3327	-0.8810
2	-0.2514	1.0318	-0.2541
3	-0.1301	1.0085	-0.1305
4	-0.1039	1.0054	-0.1041
5	-0.0518	1.0013	-0.0518
6	-0.0309	1.0005	-0.0309
7	-0.0095	1.0000	-0.0095
8	0.0000	1.0000	0.0000
9	0.0000	1.0000	0.0000
10	0.0000	1.0000	0.0000
11	3.6009	18.3295	18.3022
	Sum	33.4068	12.2509

$$BPI(T,P) = \frac{33.4068}{33.4068 + 12.2509} = 0.73$$

Set T

Set P		8	9	10	11	
		0.79	0.72	0.61	0.15	Min
0	1.00	0.21	0.28	0.39	0.85	0.21
1	0.94	0.15	0.22	0.33	0.79	0.15
2	0.90	0.11	0.18	0.29	0.75	0.11
3	1.00	0.21	0.28	0.39	0.85	0.21
4	1.00	0.21	0.28	0.39	0.85	0.21
5	0.60	0.19	0.12	0.01	0.45	0.01
6	1.00	0.21	0.28	0.39	0.85	0.21
7	0.49	0.30	0.23	0.12	0.34	0.12
	Min	0.11	0.12	0.01	0.34	

Hasudorff Distance *HD*(*T*, *P*)

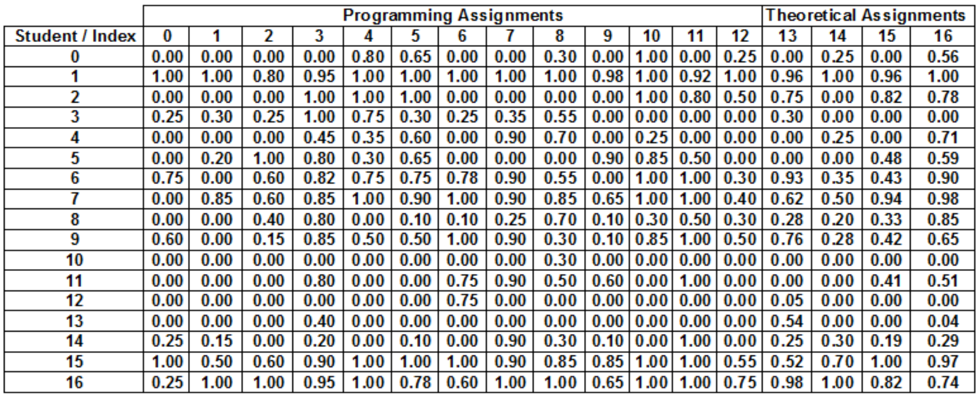
Figure 4. Computation of the Bipartivity Index (*BPI*) and Hausdorff Distance (*HD*) Metric Values for the Sample Data Set of Figure 1

(ii) Hausdorff Distance: The Hausdorff Distance (*HD*) metric [5] has been traditionally used to quantify how far are two data sets in a particular metric space. In the context of quantifying the Theoretical vs. Programming Disparity, we propose to compute the Hausdorff Distance (see below for the formulation) between the decimal percentage scores (referred to as data points) in the sets of theoretical assignments (*T*) vs. programming assignments (*P*). For every data point in data set *T* (and likewise, *P*), we determine the closest distance (in our case, the absolute difference) to a data point in the other data set *P* (*T*). The Hausdorff Distance for the *P* vs. *T* scores for a student data set is the maximum of the closest distances determined as mentioned above. The Hausdorff Distance will be thus smaller if every data point in one data set is closer to some data point in the other data set.

$$HD(T,P) = Max\left\{ \underset{i\in T}{Min}\left[\underset{j\in P}{\left| T(i) - P(j) \right|} \right], \underset{j\in P}{Min}\left[\underset{i\in T}{\left| P(j) - T(i) \right|} \right] \right\}$$

For *T* vs. *P* data sets that exhibit larger disparity, we expect several data points (assignment scores) in one data set to be appreciably different from those of the other data set. However, even if there exists one outlier data point in either of the two data sets (that is farther away from the other data points in the two data sets), the Hausdorff Distance metric has the vulnerability to get larger and not be an accurate reflection of the closeness or the extent of disparity among the assignment scores in the two sets *T* and *P*. Figure 4 displays the computation of the Hausdorff Distance metric values for the running example *P* vs. *T* data set of Section 2: we observe the presence of a lower theoretical assignment score (0.15 corresponding to index 11 in set *T*) contributes to a relatively larger *HD*(*T*, *P*) value of 0.34 (note that the next largest value among all the minimum values computed across the two data sets is 0.21).

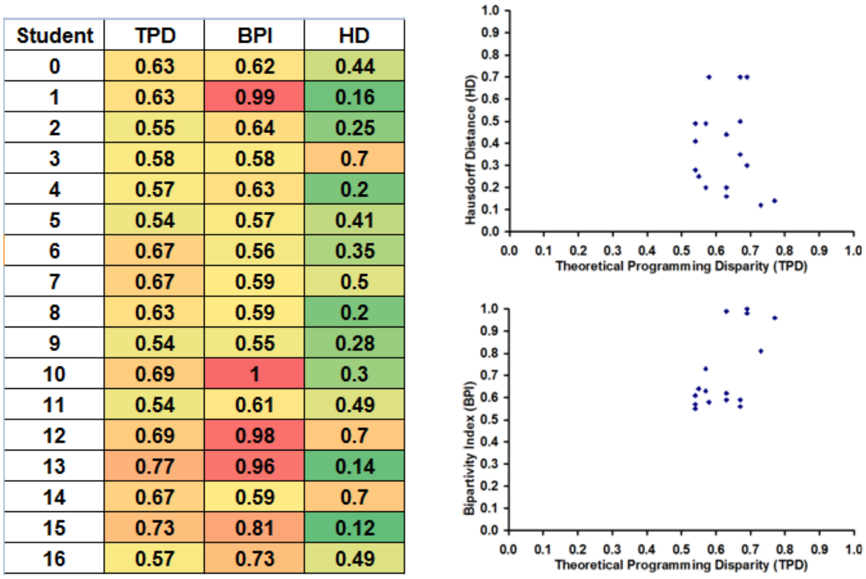
Figure 5 presents a real-time data set comprising of the assignment scores (13 programming assignments and 4 theoretical assignments) for 17 students of the CSC 228 Data Structures and Algorithms course taught in Spring 2020 at Jackson State University, MS. Figure 6 presents the values for the *TPD*, *BPI* and *HD* metrics (also visually compared using a heat map [6]) obtained for the data set of 17 students as well as plots the *TPD* vs. *BPI* and *TPD* vs. *HD* distributions. In the heat map shown in Figure 6, the red, yellow/orange and green colors are respectively indicators of high, moderate and lower values for the *TPD*, *BPI* and *HD* metrics (all of which can be represented in a scale of 0 to 1). We observe the *BPI* and *HD* metrics to have a tendency of over rating (too much red cells for the *BPI* metric) and under rating (too much green cells for the *HD* metric) the theoretical vs. programming disparity. On the other hand, the values for the *TPD* metric are within a moderate range of 0.54 to 0.77 that is sufficient enough to distinguish students with respect to the theoretical vs. programming disparity.



	Programming Assignments													Theoretical Assignments			
Student / Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0.00	0.00	0.00	0.00	0.80	0.65	0.00	0.00	0.30	0.00	1.00	0.00	0.25	0.00	0.25	0.00	0.56
1	1.00	1.00	0.80	0.95	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.92	1.00	0.96	1.00	0.96	1.00
2	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	0.80	0.50	0.75	0.00	0.82	0.78
3	0.25	0.30	0.25	1.00	0.75	0.30	0.25	0.35	0.55	0.00	0.00	0.00	0.00	0.30	0.00	0.00	0.00
4	0.00	0.00	0.00	0.45	0.35	0.60	0.00	0.90	0.70	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.71
5	0.00	0.20	1.00	0.80	0.30	0.65	0.00	0.00	0.00	0.90	0.85	0.50	0.00	0.00	0.00	0.48	0.59
6	0.75	0.00	0.60	0.82	0.75	0.75	0.78	0.90	0.55	0.00	1.00	1.00	0.30	0.93	0.35	0.43	0.90
7	0.00	0.85	0.60	0.85	1.00	0.90	1.00	0.90	0.85	0.65	1.00	1.00	0.40	0.62	0.50	0.94	0.98
8	0.00	0.00	0.40	0.80	0.00	0.10	0.10	0.25	0.70	0.10	0.30	0.50	0.30	0.28	0.20	0.33	0.85
9	0.60	0.00	0.15	0.85	0.50	0.50	1.00	0.90	0.30	0.10	0.85	1.00	0.50	0.76	0.28	0.42	0.65
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.80	0.00	0.00	0.75	0.90	0.50	0.60	0.00	1.00	0.00	0.00	0.00	0.41	0.51
12	0.00	0.00	0.00	0.00	0.00	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00
13	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.54	0.00	0.00	0.04
14	0.25	0.15	0.00	0.20	0.00	0.10	0.00	0.90	0.30	0.10	0.00	1.00	0.00	0.25	0.30	0.19	0.29
15	1.00	0.50	0.60	0.90	1.00	1.00	1.00	0.90	0.85	0.85	1.00	1.00	0.55	0.52	0.70	1.00	0.97
16	0.25	1.00	1.00	0.95	1.00	0.78	0.60	1.00	1.00	0.65	1.00	1.00	0.75	0.98	1.00	0.82	0.74

Figure 5. Real-time Data Set used to Apply and Evaluate the Proposed Approach

The tendency of the *BPI* and *HD* metrics to respectively over rate and under rate the theoretical vs. programming disparity can be observed in the case of students 10, 12 and 13, who all have just one or two submissions in one of the two categories (programming or theoretical). For such students, the majority of the entries are 0.0. The *BPI* metric tends to cluster the assignment scores for each of these students to two sets: a set of assignments for which submission has been made and a set for which no submission has been made, with the edge weights capturing the difference between these two sets. On the other hand, the *HD* metric tends to give more weight to the minimal difference in the scores between any two assignments. Both the *BPI* and *HD* metrics tend to treat all the 17 assignments as one set (i.e., one dimension) and tends to partition to two clusters (*BPI*) or find the minimum score between any two assignments; whereas, the *TPD* approach considers the problem in a two-dimensional perspective of theoretical vs. programming assignments; the clustering in the two-dimensional space gives leverage to consider four possible combinations for two clusters: the number of theoretical assignments in the sets *X* and *Y* as well as the number of programming assignments in the sets *X* and *Y* identified using spectral analysis as well as makes use of the actual number of theoretical and programming assignments to compute the Euclidean distances as formulated in Section 2. Due to the different approaches taken, we observe only a weak-moderate correlation between the *TPD* vs. *BPI* scores and the *TPD* vs. *HD* scores for the data set analyzed in Figures 5 and 6.



Student	TPD	BPI	HD
0	0.63	0.62	0.44
1	0.63	0.99	0.16
2	0.55	0.64	0.25
3	0.58	0.58	0.7
4	0.57	0.63	0.2
5	0.54	0.57	0.41
6	0.67	0.56	0.35
7	0.67	0.59	0.5
8	0.63	0.59	0.2
9	0.54	0.55	0.28
10	0.69	1	0.3
11	0.54	0.61	0.49
12	0.69	0.98	0.7
13	0.77	0.96	0.14
14	0.67	0.59	0.7
15	0.73	0.81	0.12
16	0.57	0.73	0.49

Figure 6. Comparison the Quantitative Assessments of Theory vs. Programming Disparity using the Proposed Theoretical Programming Disparity (*TPD*) Metric vs. the Bipartivity Index (*BPI*) and Hausdorff Distance (*HD*) Metrics

4. Related Work

To the best of our knowledge, we have not come across any work in the literature that focuses on assessing and quantifying the disparity found between two different categories of assignments on a per-student basis. Disparity studies in academic settings have been so far mainly focused on gender [12] and race [13] as well as on the class, as a whole (e.g., [9-11]), and not on a per-student basis. The closest work we have come across related to our topic is the work of [14] wherein the authors apply principles from phenomenography [15] and variation theory [16] to explore the practices that are needed to bridge the gap in learning Computer Science theory and learning Computer Science practice (programming). But, there are no efforts to quantify the extent of the gap (or the disparity), as is done in our paper.

Below is a review of the works that we came across in the literature that focus on studies conducted to assess the contributing factors to the success or hardship for students majoring in Computer Science and the programming component of it. In [17], the authors did a survey to find out that students think Computer Science-ability is something both innate as well as extensible through effort. In [18], the authors surveyed Computer Science (CS) student performance data for 10 years and conclude that a successful CS student needs to be strong both in critical thinking skills and the core CS skills. They observed the critical thinking skills for a CS student typically come from the Math and Physics courses and concluded that these courses need to be enforced as pre-requisites for CS courses early on in the curriculum instead of being taken along with CS courses. However, no analysis has been reported in [18] regarding the skills that influence the theory vs. programming disparity found among CS students. In [19], the authors report there is no statistically significant influence of the assessment mode (programming in a computer vs. writing a program in pen by hand) on the performance of students in a programming course. In [20], the authors observed that novice programmers tend to program using problem-solving skills obtained from domains familiar to them. In [21], the authors used reflective essays-based Attribution Theory to elucidate the internal and external causes that influence the performance of students in their first programming course.

The following works focus on analyzing the impact of one examination format on another. In [9], the authors build a model to predict the performance of students on the basis of examination formats: whether or not the performance of students in practical examinations can be predicted using their performance in the standardized examinations? The answer reported in [9] is No, as the two examination formats are observed to test different skill sets. Likewise, Haberyan [10] found no correlation between performance in weekly quizzes and examinations among students majoring in Biology. However, in [11], the authors observed that psychology undergraduates performed well in the examinations when they were also given weekly reading assignment-based quizzes throughout the course.

5. CONCLUSIONS

The high-level contribution of this paper is a spectral analysis-based approach to quantify the disparity (the proposed metric is referred to as Theoretical vs. Programming Disparity: *TPD* metric) in the scores earned by students in two categories of assignments: theoretical and programming. The uniqueness of the proposed approach is that *TPD* quantifies the disparity on a per-student basis; the typical approach in the academic community so far is to use quantitative metrics that capture the disparity for an entire class as a whole [8]. Also, traditionally for such problems, the correlation measures [3, 8, 9] are used to quantify the extent of influence on one category of assignments over another. But, to use correlation measures, we need to have the same number of assignments under both the categories. With our proposed approach, there can be a different number of assignments for the two categories. The pre-requisites are just the need to input the classification of the assignments ids (as either theoretical or programming) and the actual number of assignments under the two categories. We have demonstrated the characteristics and the uniqueness of the *TPD* metric and the spectral analysis through an exhaustive evaluation with a real-time dataset as well as through comparison with quantitative metrics that appear to be compelling enough to be suitable to capture the disparity in the student scores in the assignment categories. The disparity problem on a per-student basis among two data sets of uneven size, especially in an academic setting, has not been so far considered in the literature for quantitative evaluation; we expect the proposed *TPD* metric and its computation approach would be valuable to both academicians and researchers.

ACKNOWLEDGEMENTS

The author would like to thank everyone, just everyone!

REFERENCES

[1] Estrada, E, (2010) "Structural Patterns in Complex Networks through Spectral Analysis," *Proceedings of the 2010 Joint IAPR International Conference on Structural, Syntactic, and Statistical Pattern Recognition*, pp. 45-59. Springer-Verlag, Cesme Izmir, Turkey.

[2] Sarkar, C., and Jalan, (2018) "Spectral Properties of Complex Networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 10, 102101.

[3] Strang, G (2019) *Linear Algebra and Learning from Data*, 1st edition, Wellesley-Cambridge Press, Wellesley, MA, USA.

[4] Ernada, E., and Rodriguez-Velazquez, J. A (2005) "Spectral Measures of Bipartivity in Complex Network," *Physical Review* E, vol. 72, no. 4, 2, 046105.

[5] Birsan, T., and Tiba, D (2006) "One Hundred Years since the Introduction of the Set Distance by Dimitrie Pompeiu," *Proceedings of the IFIP Conference on System Modeling and Optimization*, vol. 199, pp. 35-39. Springer, Turin, Italy.

[6] Wilkinson, L., and Friendly, M (2009) "The History of the Cluster Heat Map," *The American Statistician*, vol. 63, no. 2, pp. 179-184.

[7] Godsil, C., and Royle, G. F (2013) *Algebraic Graph Theory*, 1st edition, Springer, Berlin, Germany.

[8] Caven, M (2019) "Quantification, Inequality, and the Contestation of School Closures in Philadelphia," *Sociology of Education*, vol. 92, no. 1, pp. 21-40.

[9] Davison, C. B., and Dustova, G (2017) "A Quantitative Assessment of Student Performance and Examination Format," *Journal of Instructional Pedagogies*, vol. 18, pp. 1-10.

[10] Haberyan, K. (2003) "Do Weekly Quizzes Improve Student Performance on General Biology Exams?" *The American Biology Teacher*, vol. 65, pp. 110-114.

[11] Johnson, B. C., and Kiviniemi, M. T (2009) "The Effect of Online Chapter Quizzes on Exam Performance in an Undergraduate Social Psychology Course," *Teaching of Psychology*, vol. 36, no. 1, pp. 33-37.

[12] Master, A., Meltzoff, A. N., and Cheryan, S (2021) "Gender Stereotypes about Interests Start Early and Cause Gender Disparities in Computer Science and Engineering," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 48, e2100030118.

[13] Kozlowski, D., Lariviere, V., Sugimoto, C. R., and Monroe-White, T (2002) "Intersectional Inequalities in Science," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 119, no. 2, e2113067119.

[14] Thune, M., and Eckerdal, A (2019) "Analysis of Students' Learning of Computer Programming in a Computer Laboratory Context," *European Journal of Engineering Education*, vol. 44, no. 5, pp. 769-786.

[15] Farton, M (1986) "Phenomenography - A Research Approach Investigating Different Understandings of Reality," *Journal of Thought*, vol. 21, no. 2, pp. 28-49.

[16] Bussey, T. J., Orgill, M., and Crippen, K. J (2013) "Variation Theory: A Theory of Learning and a Useful Theoretical Framework for Chemical Education Research," *Chemical Education Research Practice*, vol. 14, pp. 9-22.

[17] Lewis, C. M., Yasuhara, K., and Anderson, R. E (2011). "Deciding to Major in Computer Science: A Grounded Theory of Students' Self-Assessment of Ability," *Proceedings of the 7th International Workshop on Computing Education Research*, pp. 3-10, ACM, Providence, RI, USA.

[18] Yang, H., Olson, T. W., and Puder, A (2021) "Analyzing Computer Science Students' Performance Data to Identify Impactful Curricular Changes," *Proceedings of the IEEE Frontiers in Education Conference*, pp. 1-9, IEEE, Lincoln, NE, USA.

[19] Oqvist, M., and Nouri, J (2018) "Coding by Hand or on the Computer? Evaluating the Effect of Assessment Mode on Performance of Students Learning Programming," *Journal of Computers in Education*, vol. 5, pp. 199-219.

[20] Wellons, J., and Johnson, J (2011) "A Grounded Theory Analysis of Introductory Computer Science Pedagogy," *Systemics, Cybernetics and Informatics*, vol. 9, no. 6, pp. 9-14.

[21] Vivian, R., Falkner, K., and Falkner, N (2013) "Computer Science Students' Casual Attributions for Successful and Unsuccessful Outcomes in Programming Assignments," *Proceedings of the 13th Koli Calling International Conference on Computing Education Research*, pp. 125-134, ACM, Koli, Finland.

AUTHOR

Natarajan Meghanathan is a tenured Full Professor of Computer Science at Jackson State University, MS, USA. He graduated with a PhD in Computer Science from The University of Texas at Dallas in 2005. His primary areas of research interests are Network Science, Machine Learning and Cyber Security. He has authored more than 175 peer-reviewed research publications in these areas.

FORWARD CHAINING AND SELF-EMBEDDING WATERMARKING FOR TAMPER DETECTION IN A CONTINUOUS STREAM OF DATA

Sandip Hodkhasa and Huiping Guo

Department of Computer Science,
California State University, Los Angeles, California, USA

ABSTRACT

Watermarking is extensively used in various media for data transfer, content authentication and integrity. The continuous flow of data is always vulnerable to tamper. This research proposes a new watermarking scheme that detects tampering in a stream of data. The stream of data is dynamically divided into different sized groups using synchronization points. A computed watermark is embedded in each group by hashing the concatenating the current group and the next group. A secondary watermark is generated based on the current group that prevents tampering from any attacks in the current group. Watermark verification table is used to determine all possible scenarios for false results. Experiments are performed to show its efficiency. False results decrease as the group size becomes larger. Random burst attacked requires larger group size. The scheme also shows with the increase in grouping parameter 'm' which defines the synchronization point, the false positive rate decreases.

KEYWORDS

Cryptography, Digital Watermarking, Hashing, Information Hiding, Tamper Detection.

1. INTRODUCTION

Various applications that require continuous flow or streaming of data. This large flow of continuous data has applications in Internet of Things (IOT), Sensor networks and other IOT related fields [1]. Therefore, there is a constant need for data preservation, authentication, integrity and security. Research and development for security and privacy on such continuous streaming of data is exponentially growing. With new developments in various technologies, the exploitations such as vulnerabilities, complications, and loopholes in a flow of data also increases. Some of the issues include copyright infringement, data authentication, data integrity, illegal distribution, and others. [2] At the same time, data streaming over unreliable networks are subjected to data tampering and manipulation which is a concern as well [3].

Cryptography is one of the oldest forms of technology that is extremely used in data protection and security [2]. However, cryptographic algorithms can be computationally expensive because of their modular exponential multiplications and power; hence it is not widely applicable or used in Wireless Sensor Networks (WSN) and IOT [4]. Digital watermarking is a new form of technique that "complements cryptography and steganography" [5] and aids in data integrity, authentication and protects against illegal copying and tampering of data. In digital watermarking, the data is embedded into the host media (such as video, audio, images, and text)

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 147-162, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121612

just as in steganography, however, in steganography the host is embedded with a secret message whereas in watermarking the host contains several types of meta data such as its ownership, origin etc. Watermarking is computationally a lightweight solution as opposed to cryptographic algorithms that require multiple iterations of modular calculations. Hence making watermarking a great candidate for WSN, IOT and similar type of applications.

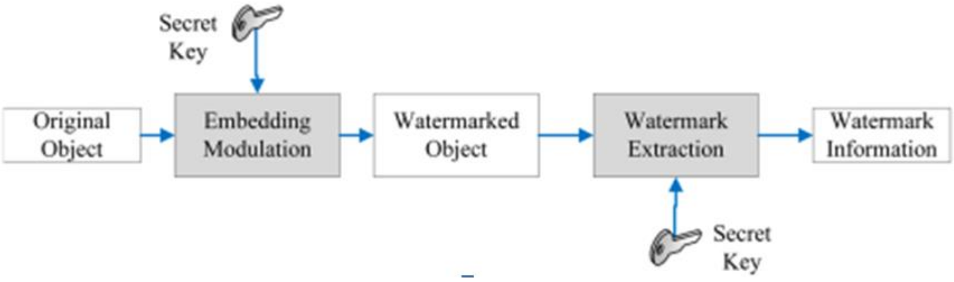


Figure 1. A watermarking technique [6]

Digital watermarking has the following major requirements [7]:

1. Transparency: A Secret data embedded into the host media that is invisible to plain sight.
2. Robustness: Watermarking should handle various attacks and cannot be easily destroyed.
3. Security: Embedded watermarking cannot be removed from the host. Removal of the watermark can lead to destroying or demolishing the host data as well.
4. Payload/Capacity: Overhead capacity needed to embed into a watermark. Embedding watermarks requires some memory.

The usage of digital watermarking provides the following [8] advantages:

1. The computation required to generate a watermark and embed into the host data is typically very lightweight. Hence it requires low consumption of energy compared to its peer technologies.
2. Since the host can embed the watermark into itself, there is low overhead for communication. This is very advantageous in WSN and IOT communications.
3. Lack of encryption and decryption algorithms, and single computation for most watermarking generation reduces end to end delay in network communication.

Digital watermarking can be categorized into the following branches [9] [10] [11]:

1. Fragile: Watermarking that is extremely sensitive to modifications and can sense slightest change in the watermarked data. It is useful for author authentication of digital content.
2. Robust: Such watermarking can endure various voluntary or involuntary attacks that can cause data manipulation. Attacks on host media carrying robust watermarking typically degrades the host media. Therefore, robust watermarking is extremely useful in illicit copying of media.
3. Semi-Fragile: These types of markings make use of both the above-mentioned schemes. Semi-Fragile watermarking can locate the location of tampering as well as the algorithm can restore the watermark if tampered.

This current research makes use for fragile watermarking to detect tamper in a continuous stream of incoming data. This source of incoming data can come by myriad applications from online websites to sensor network such as stock market to various types of sensors installed in smart cities and forest for temperature detection. In this research we assume there is continuous flow of data. This technique is applicable to other types of data, however, this research uses only integers for simplification.

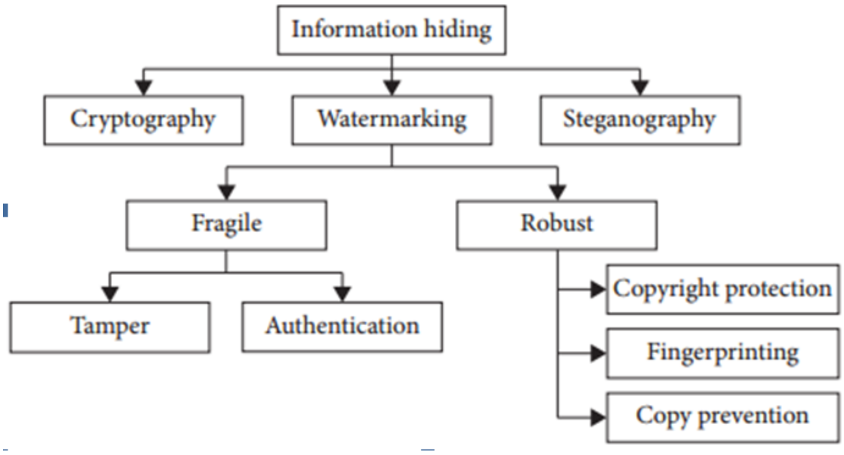


Figure 2. Common information hiding techniques [5].

In chapter 2, previous related works are reviewed and their flaws, whereas in chapter 3, the proposed algorithm is discussed. Primary and secondary watermark is introduced and the process of embedding, and extraction of the watermarks is also discussed. Table 1 shows the water detection table and Table 2 summarizes the list of all symbols used. Chapter 4 talks about the generation and the use of synthetic data. It also mentioned about the false positive and false negatives values that can get generated using this approach. Watermarking detection rationale is also introduced. The watermark rationale table discusses how false positive and negative values are calculated. Also, pros and cons of using the two watermarking scheme is considered. Different types of attacks are conducted, and the results of such attacks are examined and reviewed to deduce conclusions. The conclusion is discussed in chapter 5.

2. RELATED WORK

The communication using text and numbers requires least amount of bandwidth among other modes of communications and yet can be one of the most complicated media to be watermark. [12] There are various techniques to watermark text and integers. This current research work is a very special type of text communication: a continuous flow of streaming data that is comprised of integers. Nonetheless, the algorithm mentioned in this research can be effortlessly revised to watermark text and other media. There are few methods that can be applied in watermarking such as [11] spatial domain techniques, and frequency domain technique. Spatial domains consist of least significant bit (LSB) coding which is the technique used in this research.

It is perceived that that Guo et al. [13] were the first to use fragile watermarking scheme to embed watermarking in a continuous flow of data. Their scheme chains group of inflowing data and embed watermark into it. In their scheme they propose a technique of dynamically grouping data which ends with a synchronization point which is a specific data element. This data element is defined by a hash computation of the data element whose value mod ‘m’ is zero, where ‘m’ is a number kept secret and only known to source node and sink node or end clients. Hash values are calculated using the secure hash function of each element in the group and then of the groups. Then a hash function is again performed of two consecutive groups. The hash values are then used to embed into the data.

A lightweight chained watermarking (LWC) scheme was introduced [14], where they improved upon [13] by performed less regressive of hash computation and tried address few issues by improving on computational overhead. In [4] FWC-D scheme was proposed. The authors added a group delimiter instead of grouping dynamically based on synchronization point.

[15] used a dual-marking fragile watermarking system based on the character. They also dynamically grouped data and embedded chained watermark. In their algorithm they used blank spaces as the watermark which generated from the characters that were converted from numeric data type.

[14] does not account for any false positive or false negative values. In [15] false positive rate is still too high and there seems to be no evaluation for the false positive rate with group size or for false negative rate. In [13] there is a trade-off between the security and precision. Most of the above-mentioned work does not account for small sized group and have high embedding and extraction time. It seems imminent that there is a requirement for a new algorithm that can provided integrity and remain tamper proof for all data elements within a group.

3. Proposed Algorithm

This proposed algorithm uses fragile watermarking system where watermark is embedded into the dataset, the dataset will experience some sort of minor distortion. This work also employs watermarking in the least significant bit scheme. The embedded algorithm is simple to implement and able to handle large sets of data in relatively short time.

3.1. Grouping and Synchronization Point

Assume that there is an endless and continuous flow of data. After the data is sensed by a sensor, this algorithm should be applied and then send to the receiver. The sender will have to form two groups and store data into the two buffers. Each incoming data element is filled into a group, the size of which is determined by synchronization point using following equation:

$$H i \bmod m == 0 \qquad \text{Equation 1}$$

Where H is a cryptographic hash function such as MD5 and SHA for each element s_i. Such hash function takes a variable length string as an input with a key and returns a fixed length hash value of the string. As each incoming data element goes through a hash function, equation 1 is performed to check if a data element is a synchronization point. Once the synchronization point has been encountered, that data element marks the end of that current group, and the buffer ends there forming one group. Similarly, second group is formed using the hash function as mentioned above and the list of data elements are saved into a second buffer. A set S is defined by a list of elements $s_{i,}$ such as $S = \{s1, s2, s3, s4...\}$. Correspondingly, H_i denoted the hash value of each element s_i. It is to be noted that a synchronization point is dependent on the value of 'm'.

Another factor that drives the size of the group is 'L' defined as lower bound of a group size. Lower bound is the minimum size of the group that is required to form a group. Even though a synchronization point is encountered, if the size of the group is lower than 'L' then the algorithm continues to find another synchronization point. As the two groups are formed. It is later analyzed how the value of 'L' is critical to the authentication of the data at the receiver's end. However, as the range of the group grows up to 'U' - upper bound, (or the maximum value the group size can be) then the group ends.

Algorithm 1: Creating List of Data

```
Create List1, List2
List1 = dataGenerator (int $m$, $L$, $U$)
List2 = dataGenerator (int $m$, $L$, $U$)
Function: dataGenerator (int $m$, $L$, $U$) → List list
while (True):
    number = generate a random number between a range
    subList.append(number)
    hashValue = $H$(number)
    return subList if(hashValue % $m$ == 0) and size(subList) > $L$
    return subList if(size(subList) == U)
```

3.2. Watermark Generation and Embedding

Once the two list is created, they are identified as two groups: current and next group, as *currentGroup* and *nextGroup*. In the *currentGroup* all the data elements of the *List1* are concatenated., whereas, in the *nextGroup* all the data elements of the *List2* are concatenated. The data elements in *List2* are the data elements that follows data elements in *List1*. Now two watermarks are generated involving the two groups as follows in Algorithm 2 and 3. As the watermark is embedded that data is then sent to receiver.

Algorithm 2: Create Group Hash

```
Function: CreateGroupHash(List: list1, list2) → List: Hash1, Hash2
    currentGroup = {$s_1||s_2||s_3\ldots$}
    nextGroup = {$s_n||s_{n+1}\ldots$}
    //When getting hash values the last two bit of each data element is ignored
    Hash1 = $H$(currentGroup||nextGroup ||key1)
    Hash2 = $H$(currentGroup||key2)
return Hash1, Hash2
```

Algorithm 3: Watermark Embed

```
Function: WatermarkEmbed(Hash1, Hash2)
    for each data element ‘i’ in List1:
        replace last bit of ‘i’ by last bit of Hash1
        replace second last bit of ‘i’ by last bit of Hash2
        move right to the next element in List1
        move left to the next bit in Hash1 and Hash2
    end for
```

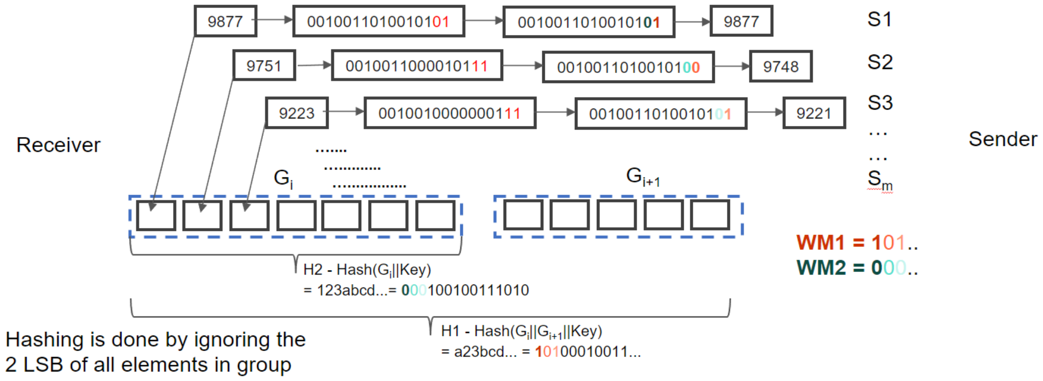


Figure 3. A graphical representation of the embedding process

3.3. Watermark Detection

As the data is received from the sender at the receiving end, the watermark extraction process is implemented. The extraction and the embedding processes are very similar. As the data is being obtained by the receiver, sink or the server, the incoming should be added to the buffer or list as it was done in *List1* and *List2* algorithm 1. Two sets of lists are created that terminates at the synchronization point using equation 1. Just to note that these data already contain the two watermarks. Once the lists are formed, algorithm 2 can be used again to generate the hash values of the two lists. Also point to be noted that when the hash values are generated, the algorithm ignores the last two bit of each data elements and then concatenates each data element to the group. Once the hash values are generated, two watermarks are created as *WMG1* and *WMG2*. *WMG1* which is the primary watermark is generated from *Hash1* by concatenating the extreme right 'k' bits ('k' being the size of the current most group or list of data received). Similarly, like *WMG1*, *WMG2* can be generated from *Hash2* which is referred to as secondary watermark.

Correspondingly, using the new lists that are created at the receiving end, *receiveList1* and *receiveList2*, the embedded watermark is extracted. This extracted watermark is then compared to watermark that is generated using the data received at the receiving end mentioned in the previous paragraph. The watermark that is extracted by extracting the last bit of every data element (which is also the watermark generated by getting the hash value of *H* (*currentGroup*||*nextGroup* ||*key1*)) is referred to as the primary watermark or *WME1*. The watermark that is extracted by extracting the second last bit of every data element (which is also the watermark generated by getting the hash value of *H* (*currentGroup*|||*key2*)) is referred to as the secondary watermark or *WME2*.

With the availability of two watermarks, the verification for integrity and modification makes it less complicated and easy. We use *WME1* to check the integrity of the data, it checks if a group of elements have been deleted, while *WME2* is used to check if there is a modification in a group. Two types of verification will be done using *WME1*: preliminary and final verification. Preliminary verification is done using the *WME1* for the current and next group, while final verification is based on the preliminary and final verification of previous group and current group using *WME1*. *WME2* will be used when the final verification of the current group is turned out to be false or the *WME1* and *WMG1* are not matched.

3.4. Watermark Verification

To verify the integrity of the group of incoming data, two buffers or lists of data are constructed as mentioned above. Since there are two groups of data, we have two types of verification: preliminary and final denoted as pV and V respectively, and two watermarks are present. The verification can be done using algorithm 4 as shown below. The preliminary and final verification of the previous group is assigned as $pV0$ = False and $V0$ = False in the beginning. After each iteration, the value of $pV1$ and $V1$ is assigned to $pV0$ and $V0$, where $pV1$ and $V1$ are the preliminary and final verification of the current group. $V2$ is the verification result of comparing the secondary watermark embedded and generated at the receiving end.

Algorithm 4: Watermark Verification

```
V2 = True if (WME2 == WMG2) else V2 = False
if (WME1 != WMG1):
    pV1 = False
    V1 = V0 & pV0
else: V1 = pV1 = True
```

This current work makes some assumptions and takes leniency in the data. When the watermark is being embedded and extracted, each data element of group set goes through a small distortion i.e. the last two bit of each data element is ignored while generating watermark, and the watermark once generated is embedded into the last two bit of each data element. Watermark one can also we called as the integrity watermark and watermark two is also called the anti-modification watermark. The watermark detection and verification are based on the following Table 1. For cases 1-4, where the watermark is matched it means that the group that is modified or if a group is missing, that has been successfully verified. Rest of the cases are discussed below in Table 1.

Table 1. Watermark verification table

Predicates	**Cases**	**Previous Group**		**Current Group**		**Current Group**
		Group 1 - G1		Group 2 - G2		
		PV0 (WM1)	V0 (WM1)	PV1 (WM1)	V1 (WM1)	V2 (WM2)
Watermark Match	1	TRUE	TRUE	TRUE	TRUE	
Yes	2	FALSE	FALSE	TRUE	TRUE	
		Entire group between previous and the current group may be absent				
	3	FALSE	TRUE	TRUE	TRUE	
		Entire group between previous and the current group are absent				
Watermark Match	4	TRUE	TRUE	FALSE	TRUE	
No	5a	FALSE	TRUE	FALSE	FALSE	FALSE
		G2 Modified				
	5b	FALSE	TRUE	FALSE	FALSE	TRUE
		Initial False Positive, but WM2 confirms group missing between G1 and G2				
	6a	FALSE	FALSE	FALSE	FALSE	FALSE
		G2 Modified				
	6b	FALSE	FALSE	FALSE	FALSE	TRUE
		G1 and G3 modified, Groups missing				

Case 5a: In case 5a, if the preliminary verification and the final verification of the current group is false, it means that either the current group or the next groups is modified, i.e., WM1 did not match. Back checking with the previous group it turns out that the previous group is verified to be true. Therefore, at this point, WM2 is used to verify the authenticity of the current group. If the WM2 is false, that only means the current group is modified.

Case 5b: In case 5b, if the preliminary verification and the final verification of the current group is false, it means that either the current group or the next groups is modified, i.e. WM1 did not match. Back checking with the previous group it turns out that the previous group is verified to be true. Therefore, at this point, WM2 is used to verify the authenticity of the current group. If the WM2 is true, that only means the current group was not modified. This results in conclusion that there must be a group missing between G1 and G2

Case 6a: In case 6a, if the preliminary verification and the final verification of the current group is false, it means that either the current group or the next groups is modified, i.e., WM1 did not match. Back checking with the previous group, if it turns out that the previous group is verified to be false, then it means that at this point, either there is a group missing between the previous and the current group or there is a group missing between the current group and next group. WM2 is used to verify the authenticity of the current group. Since the WM2 is false that only means the current group is modified.

Table 2. List of Symbols and its meaning

Symbol	**Meaning**
H	Cryptographic Hash Function
s_i	Data Element
S	Data Set Element
m	Synchronization Point
L	Lower bound of group size
U	Upper bound of group size
List1	all the data elements of *currentGroup*
List2	all the data elements of *nextGroup*
WMG1	the primary watermark is generated from Hash1
WMG2	the secondary watermark is generated from Hash2
WME1	primary watermark extracted from the received list
WME2	secondary watermark extracted from the received list
PV0	Preliminary Verification for previous group using WM1
V0	Final Verification for previous group using WM1
PV1	Preliminary Verification for current group using WM1
V1	Final Verification for current group using WM1
V2	Verification for Watermark 2
WM1	Watermark 1, based on the grouping of current and next group
WM2	Watermark 2, based on the current group only

Case 6b: In case 6b, if the preliminary verification and the final verification of the current group is false, it means that either the current group or the next groups is modified, i.e., WM1 did not match. Back checking with the previous group, if it turns out that the previous group is verified to be false, then it means that at this point, either there is a group missing between the previous and the current group or there is a group missing between the current group and next group. WM2 is used to verify the authenticity of the current group. Since the WM2 is true that means that current group is not modified, which means that either the previous and next group is modified or there are groups missing.

4. EXPERIMENTS

Synthetic data is used for a controlled experiment, an infinite flow of data, *S* is generated using a standard random function generator. Data ranging from 9000 to 9999 is generated as integers. Data is uniformly distributed between the range. The key values such as group separator - *m*, the two key *k1*, and *k2*, lower bound of the group *L*, are decided and assumed to be kept secret. The algorithm is tested with a minimum of 10000 data elements. The system specifications for the simulation are Windows 10 Pro, Intel(R) Xeon(R) CPU E5-1620 0 @ 3.60GHz.

4.1. Attack

The watermarking model was tested by various types of attack. The attacks were also randomly generated based on the Poisson's distribution. The random Poisson's distribution function was implemented, and it generates a random number. Based on the number generated, the following attacks were performed using the modulus of the iterations:

```
If (iteration % x == 0):
        "Modifying by appending"
Else if (iteration % x == 1):
        "Modifying by Deleting"
Else if (iteration % x == 2):
        "Modifying by Modification"
Else if (iteration % x == 3):
        "Modifying by elimination a group"
```

Where '*x'* is the random value that is generated using the Poisson's distribution. This random value '*x*' is an average of a list of ten random numbers and then used in the experiment. The *iteration* is a counter that runs and counts the random data element that is being generated. The attacks that are performed are based on the *iteration* and *x* as following:

Modifying by Appending: As the groups are formed, the groups are modified by adding data elements into the group. Random data elements were added into the current group post watermarking is embedded. After modification, the group is then sent to the receiver.

Modifying by Deleting: Once the group is formed and watermarking is performed into the data set, a data element is randomly picked from the current group and deleted from the group reducing the size of the group. The modified group is then sent to the receiver's end for verification.

Modifying by Modification: In this case, modification is done by taking the data set and picking a random element within the current watermarked data set and modifying that value of the data element by some type of operations. In this case we modified a data element by adding four to the integer. Issues with modifying the data element by a value less than four is discussed later.

Modifying by Elimination in a group: In this situation of modification, the modification is done to the entire group. As the data set or group is watermarked, after being watermarked, the entire watermarked group is substituted but another set of random data elements. This new set of random data elements may or may not be of the same size.

4.2. False Values

4.2.1. False Negative

False negative is when the data is modified after watermarking however, the algorithm for watermark detection doesn't recognize the data tampering or data modification. This experiment uses different values of lower bound groups and later results are discussed based on different sizes of the groups.

4.2.2. False Positive

False positive is when the data is not modified after watermarking, however, the algorithm for watermarking detection recognizes the data set as being tampered or modified. This experiment also accounts for data showing false positive for different sizes of lower bound groups. The following sections discusses different scenarios and reasons where false flags are raised.

4.3. Least Two-Bit Modification

Least two-bit modification of the data is an important modification which dictates the verification of the watermark that was embedded. When making a modification in the last two bit of the data, the watermarking is performed by ignoring the last two bit of the group data, when the data is watermarked, and if there is an attack where the attacker adds a value from one to three, such a modification remains un-noticed. The following example gives a better idea.

Hash is calculated ignoring the last two bit.
Ex: 9726 becomes 9724, when ignoring the last two bit.
Therefore, $H(9724) = H(9725) = H(9726) = H(9727)$ Equation 2

If after watermarking, there is a modification made by the attacker which only changes the least two bits, then watermarking verification will remain undetected. Hence, this algorithm can be used only where slight modification is acceptable.

4.4. Matching Extracted Watermark

There are instances where the post modification in a group of data set, or modification in the entire group of data still results in same watermark. Since the watermark is same, this results in false negative, where the data is modified but the algorithm is not detected. This instance typically occurs for low-sized groups. The subsequent is an example of such a case, the size of the group here is 5:

DATA FROM SENDER

Original Data:	[9759, 9785, 9738, 9040, 9776]	
Group + Key:	97569784973690409776 ‖ 11	(Ignoring the last 2 bit)
Hash Value:	7468bd11bcf7572a0066ec78efc139ca	
Embedded Data List:	[9757, 9786, 9738, 9042, 9777]	
Embedded Data List:	[9761, 9786, 9738, 9042, 9777]	(After Attack)
Hash Value in bits:	74…. = 0111 0100 …	
Watermark to Embed:	011101	

DATA TO RECEIVER

Received Data:	[9761, 9786, 9738, 9042, 9777]	
Group + Key:	97609784973690409776 ‖ 11	(Ignoring the last 2 bit)
Hash Value:	75e1e7e7789b10398d3a91152edc4876	
Embedded Data List:	[9757, 9786, 9738, 9042, 9777]	
Embedded Data List:	[9761, 9786, 9738, 9042, 9777]	(After Attack)
Hash Value in bits:	75…. = 0111 0101 …	
Watermark to extract:	011101	

Even though the data set has been modified, the watermark that is inserted and extracted are same. It can be argued that because the hash function used in this algorithm is MD5, and MD5 is not a great security hash function, there are hash valves which have similar values. However, the next section tackles other secure hash functions. It is important to note that, for small groups, there can be several false negatives.

4.5. Lower Bound Group Size

Since, in the previous section it was determined that the low-sized group plays a crucial role in determining the rate of false negative or positive values, the following graphs show the decreasing of the false values as the group size increases. It can be seen in the Figure 4, as the group size increases the false negative and positive values decreases. Two hash functions were used, MD5 and SHA256. The hash values obtained were then used as forward and backward embedding.

In forward embedding, the bits from the beginning of the hash values are used for watermarking, whereas in backward embedding, the trailing bits are used for watermarking the current group of data. Using both methods, with the increase in group size, the false results decreased. The false positive and false negatives both converge to zero at the same time as the lower limit of the group size becomes 10. In Figure 5, there is a direct comparison of the two functions shown with forward and backward embedding. There is typically not a major difference. They both follow a similar path and pattern, with bottoming out at the lower bound of group size of 10.

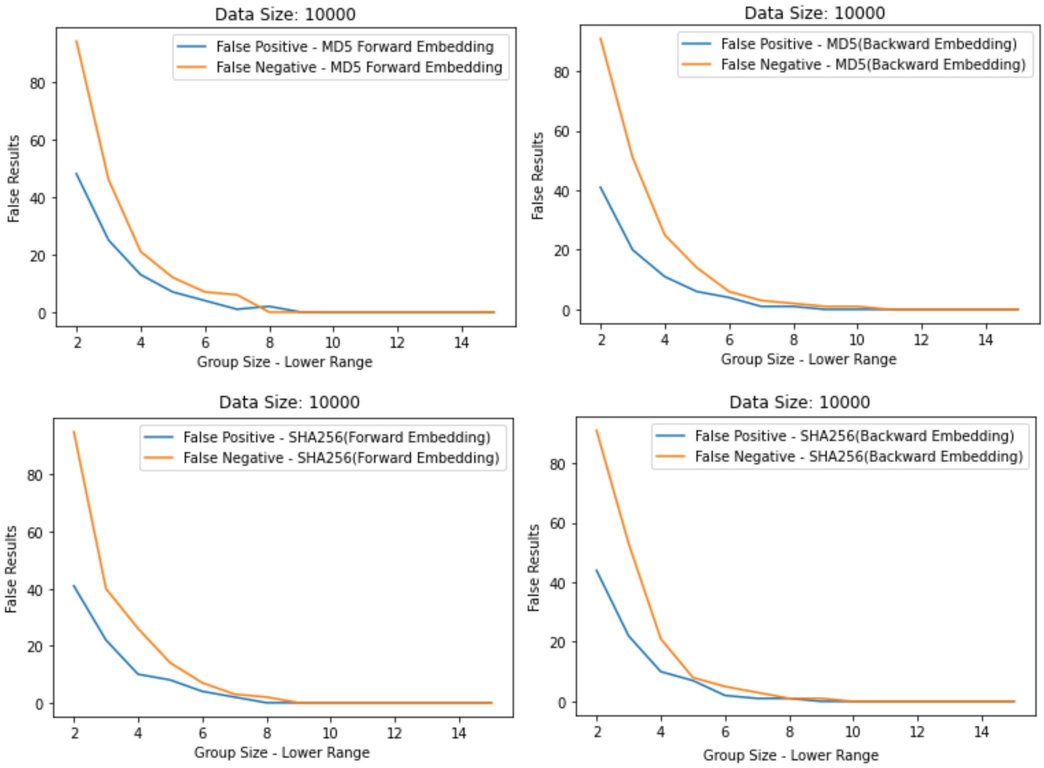


Figure 4. MD5 and SHA265 false positive and false negative values versus lower bound group size for a data size of 10000

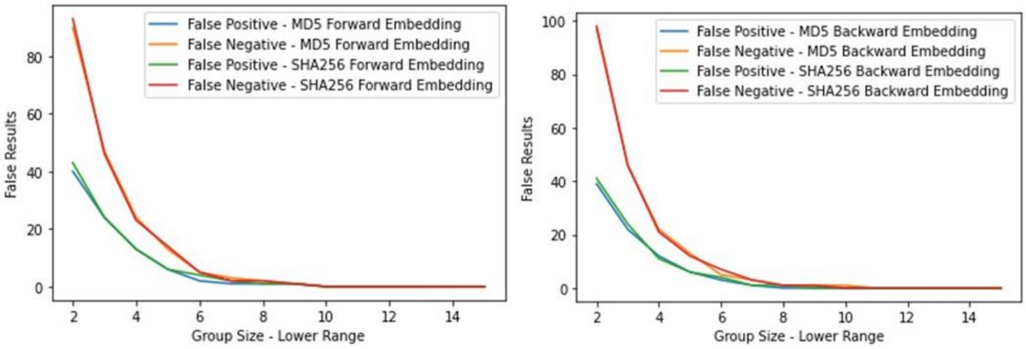


Figure 5. MD5 and SHA256, Forward Embedding and Backward Embedding comparison

4.6. Watermark Scalability

Figure 6 shows the watermark scalability. Figure 6 shows as the data size increases, the time required for watermark embedding and extraction also increases. It can also be seen that the embedding and extraction time frames are different for different group sizes. While a group size of 10-20 takes the least amount of time, a group size of 50-60 takes the maximum amount of time. The time calculated in Figure 6 is the time required to embed plus the time required to extract the watermark. Figure 6 shows the scalability of the size of data with respect to the size of group and the false results. As it can be seen, with high data sizes the false result increases, however they decay as the lower bound of the group size increases. It can be clearly seen that regardless of the data size, all of plot lines converges to zero false results when the group size approaches to 10.

4.7. Burst Attack

Burst attack was performed in this model in which a modification is repeated ‘μ‘times also known as the burst attack length. In this attack, insertion, deletion was done at equal probability and applied randomly using Possion’s distribution function. The attack is done multiple times within the same group and pattern has been observed. The attack was performed with:

1. Random Size ‘μ‘ - a random number was generated and assigned to μ each time and the insertion or deletion was done μ times.
2. Fixed Size ‘μ‘ – The insertion or deletion was done for four iterations, and in each iterations the attack was incremented by one.

As it can be seen in Figure 7, the false positive rate (false positive over data unaffected) falls as the group size increases. Small groups tend to have high false positive rates as small groups have high false positive. The false positive rate decreases with increase in group size which also means the false positive rate decreases as well. It seems that when that attack is random, the algorithm requires a high lower bound compared to the previous attacks that was performed.

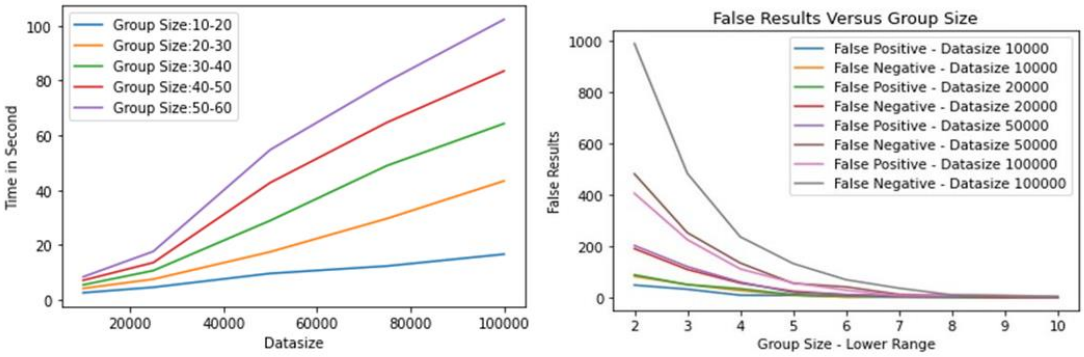


Figure 6. Scalability of watermarking shown with different group size and data size.

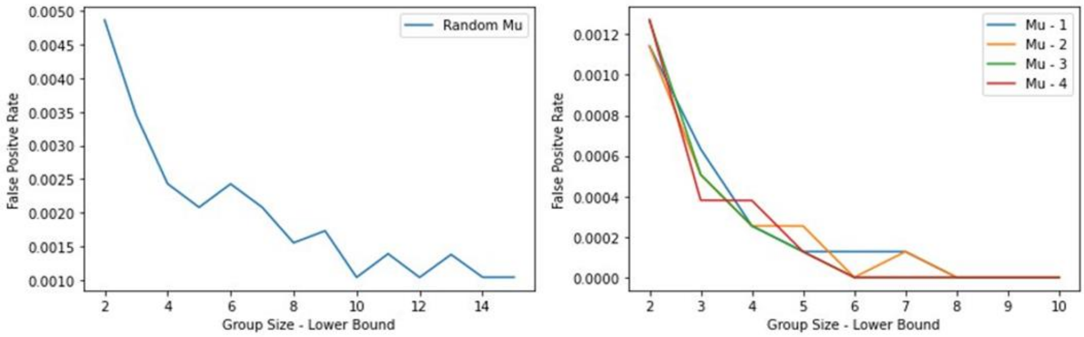


Figure 7. False positive rate versus lower bound of group size with µ varied randomly and with µ increasing from 1 to 4

4.8. Group Attack

The final attack was performed to measure the false negative rates when incoming group or groups were deleted. As the incoming groups are forming, the attack deletes the next group(s) that are being formed. The number of groups to be deleted depends on random number based on the Poisson's distribution. It was observed that the false negative values were zero and the algorithm detected all the modifications that were being made during the inflow of the data.

4.9. Grouping Parameter – 'm'

In this watermarking scheme, 'm' defines the synchronization point. which is used as a parameter to form groups. The size of the group partially depends on 'm'. The algorithm ensures that the size of the group is maintained within the range of the group that is bounded by lower and upper bound. Figure 8 (Left image) is taken from [15] where they compare the group parameter and the false rate. Whereas, in Figure 8 (Right and Middle image) it can also be seen that the false rate decreases as the group parameter increases. In [15], the lower bound 'L' is 50, whereas in this experiment the value for 'L' is very low for which the false rate tends to zero. Figure 8 (middle and right image) shows two images for the same parameters – group parameter m and false positive rates, however the group range is different. For the image in middle the group range is 4 and the right plot the range is 8. Also, it must be noted that for small group sizes the false positive rate tends to zero as the group parameter 'm' increases.

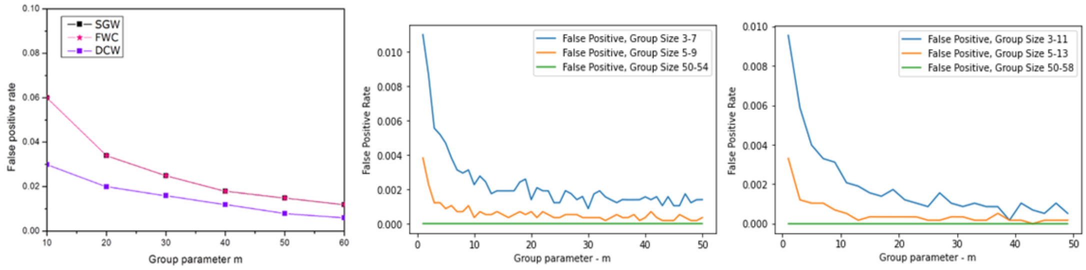


Figure 8. (Left) Results taken from [15] which compares false positive rate with group parameter, (Middle and Right) false positive versus group parameter for different range of lower and upper bound

5. CONCLUSION

As it can be seen, this algorithm has established that with a minimum of a group size of 10, data elements have successfully shown zero false results. When groups are sized under ten, false positive and false negative results are generated. Therefore, an application will have to alter data the group size to this threshold size in order to maintain the data integrity Another consideration to be noted that the application also has be adjusted so as the application can tolerate small distortion of data. Because this algorithm embeds the watermark by modifying that last two bit of data, the data is slightly modified. Also, it needs to be noted that when modifying the data i.e., tampering a data element, altering the value by less than three will make the data undetected. In other words, the alteration of the data elements needs to be more than three for the algorithm to detect any tamper. However, if the sensor can tweak the data in such a way that the least two significant bits does not impact its use in that application, then the algorithm can successfully maintain its integrity as well as prevent various types of attack. The current model used two different types of hash functions, and when the watermark was forward embedded and backward embedded, it produced similar results. They emit same false positive and false negative and converge to zero false results around the same lower bound of group size. In this research, the algorithm successfully mitigates false positives and negatives while maintaining the integrity of the data. The experiments successfully link the data that are formed into groups and watermarks can be embedded into the groups with small distortion. With slight modifications, the algorithm can be used successfully in various applications.

REFERENCES

[1] X. Sun, J. Su, B. Wang and Q. Liu, "Digital Watermarking Method for Data Integrity Protection in Wireless Sensor Networks," International Journal of Security and Its Application, vol. 7, no. 4, pp. 407-16, 2013.

[2] I. Kamel, O. Al Koky and A. Al Dakkak, "Distortion-Free Watermarking Scheme for Wireless Sensor Networks," in International Conference on Intelligent Networking and Collaborative Systems, Barcelona, Spain, 2009.

[3] B. Wang, X. Sun, Z. Ruan and H. Ren, "Multi-mark- Multiple Watermarking Method for Privacy Data Protection in Wireless Sensor Networks," Information Technology Journal, vol. 10, no. 4, pp. 833-40, 2011.

[4] I. Kamel and H. Juma, "Simplified watermarking scheme for sensor networks," International Journal of Internet Protocal Technology, vol. 5, no. 1-2, pp. 101-11, 2010.

[5] U. Khadam, M. Iqbal, M. Alruily, M. Al Ghamdi, M. Ramzan and S. Almotiri, "Text Data Security and Privacy in the Internet of Things: Threats, Challenges, and Future Directions," Wireless Communications and Mobile Computing, vol. 2020, pp. 1-15, 2020.

[6] R. Wazirali, R. Ahmad, A. Al-Amayreh, M. Al-Madi and A. Khalifeh, “Secure Watermarking Schemes and Their Approaches in the IoT Technology: An Overview,” Electronics (Basel), vol. 10, no. 14, pp. 1744-1772, 2021.

[7] R. Saxena, N. Tiwari and M. K. Ramaiya, “A Survey Work on Digital Watermarking,” International Journal of Latest Technology in Engineering, Management & Applied Science, vol. 4, no. 5, pp. 35-37, 2015.

[8] X. Yu, C. Wang and X. Zhou, “Review on Semi-Fragile Watermarking Algorithms for Content Authentication of Digital Images,” Future Internet, vol. 9, no. 4, pp. 56-73, 2017.

[9] J. Park, S. Jeong and C. Kim, “Robust and Fragile Watermarking Techniques for Documents Using Bi-directional Diagonal Profiles,” Information and Communications Security, vol. 2229, pp. 483-494, 2001.

[10] G. Zhang, L. Kou, L. Zhang, C. Liu, Q. Da and J. Sun, “A New Digital Watermarking Method for Data Integrity Protection in the Perception Layer of IoT,” Security and Communication Networks, vol. 2017, pp. 1-12, 2017.

[11] P. Singh and R. S. Chadha, “A Survey of Digital Watermarking Techniques, Application and Attacks,” International Journal of Engineering and Innovative Technology, vol. 2, no. 9, pp. 165-175, 2013.

[12] Y. Li, H. Guo and S. Jajodia, “Tamper detection and localization for categorical data using fragile watermarks,” in Association for Computing Machinery, New York, 2004.

[13] H. Guo, Y. Li and S. Jajodia, “Chaining watermarks for detecting malicious modifications to streaming data,” Information sciences, vol. 177, no. 1, pp. 281-98, 2007.

[14] I. Kamel and H. Juma, “A lightweight data integrity scheme for sensor networks,” Sensors, vol. 11, no. 4, pp. 4118-36, 2011.

[15] B. Wang, W. Kong, W. Li and N. N. Xiong, “A Dual-Chaining Watermark Scheme for Data Integrity Protection in Internet of Things,” Computer, Materials & Continua, vol. 58, no. 3, pp. 679-95, 2019.

A STOCK FORECASTING APP USING MACHINE LEARNING AND BIG DATA ANALYSIS TO HELP GUIDE DECISION MAKING WHEN INVESTING IN THE STOCK MARKET

Alex Xie[1] and Yu Sun[2]

[1]Chadwick School, 26800 S Academy Dr, Palos Verdes Peninsula, CA 90274
[2]California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

ABSTRACT

Currently, there is increasing participation in investment, especially in the stock market, as it appeals to more average citizens. One common roadblock these individuals receive is the lack of information about the economy, politics, regulation, etc., which could all affect the stock market. This app addresses this problem by collecting mass information from third-party social media. Information from social media platforms has its advantage mainly due to the citizen involvement and expertise some users may resemble. Pulling large amounts of data from these social media users avoids bias and establishes reliable sourcing. Using this information as a predictor, the app computes data and effectively predicts the performance of the stock for the next three days. By doing this, users of this app could easily get informed through instant quantitative prediction instead of having to read all over social media. This app also indirectly manages people's wealth because it allows users to make smarter decisions, thus increasing their money potential. In certain areas, this app is able to perform the same duty as a licensed wealth manager.

KEYWORDS

Stock, Asset, Investment, Exchange.

1. INTRODUCTION

The benefits of investment are enormous to both the investor and the company. Business investment often helps small businesses to grow and leads them out of budget hardship. Money that a company receives stimulates the expansion of projects and the efficiency of reaching its goals. Attracting investors also means the company is trusted by others and thus will grow in popularity in the eyes of the citizens [1]. On the other hand, the benefits to investors should not be neglected. Investors hope for long-term returns that add to their income. Investing in well performing companies is the quickest way to multiple assets without having to work physically. Besides, investment also outperforms inflation in the economy, as the return rate is usually higher than the rate of inflation [2].

Since more people are willing to give their money for investing, stock trading platforms are innovating to provide the most convenient experiences to customers in order to attract more users. As the stock market is growing and more people become involved in investing in stocks hoping to enlarge their wealth, many barriers start to unveil due to the lack of expertise and

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 163-175, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121613

background knowledge of certain companies' performance. There are many unpredicted fluctuations in the stock market. People lose money because they were not able to obtain or be alerted quickly enough to act [3]. Additionally, the reason why more average citizens are transitioning into digital investments is that it appears to be the easiest method to earn money. This illusion lures people into spending their money on investments. In other words, unprofessional traders bet their money and underestimate the treacherous market.

First off, in terms of trading platforms, there are numerous apps and websites where users can legally exchange stocks in the U.S. Inside of those stock trading platforms, the developer usually implements different forums where users and professionals can discuss and share trends and performance of certain stocks. There are also new feed sections where users can be informed of the latest news regarding the current market in the economy. These are ways that help advise users on their trading decisions in order to improve their user experience [4]. However, this method does have its flaw. First of all, time is a huge constraint for users reading news and forums. Spending a long time reading these suggestions and information does not provide users with the quickest method of obtaining knowledge. Plus, there are many forums out there with different descriptors of the companies, choosing the right forum can be super time-consuming. On the contrary, it ruins many users' experience in investing. The second flaw in these forums and news feeds is the lack of comprehension. Keep in mind that these users are often inexperienced with many terminologies that more knowledgeable people use when trading. Even if these average users devote endless hours to reading and researching, they might not be able to understand the general meaning of these texts. Another serious problem with these forums is that it is public to everyone, and the information everyone is sharing could be either fraud or unreliable [5]. Users often only experience a small pool of information from a designated source. The problem with this is that since people are not accessing a wide range of sources, they might often get misled by these biased and subjective voices. There are many cases where institutional investors did a rug-pull, causing retail investors to lose money [6]. All of these are serious problems in the trading system right now, and this stock forecasting app aims to solve these issues.

Our goal in developing this app is to ensure all prospective traders, no matter what prior experience they resemble, are able to be better informed on their investing decision-making. We want to open out the gate and let all information into the hands of the users efficiently and effectively. This is not a stock trading app, it is simply an app where users can track stocks and see predictions of the stock's performance. When we built this app, we considered many features worth adding to improve the functionality of the app. The first and most important feature is pulling social trends from third-party social media and compiling them for analysis. This is a very crucial step of the development because we need to make sure we have access to all information out there in order to make our predictions more inclusive and reliable. The second feature we included comes after compiling the information, which is to use numbers to make predictions of the stock's performance in the next three trading days. Displaying the percentage increase or decrease of the stock to inform people effectively. Last but not least, we have features that focus on user experiences, such as a search to all 6000 stocks, a watch list, and a daily trend that helps pertain to the user's needs.

Determining whether the prediction is accurate was what we mainly focused on while building this app. Since we are using a machine learning model, there needs to be large amounts of data and various trials to formulate a representative machine. Choosing a display model of the graph is also important to our design. Whether to use linear or polynomial expression affected the outcome. We tested these graph algorithms and found out that each model delivers a different result, even though the data and information inputted are identical. We also ran many algorithms when testing our data and prediction. We started off by only implementing a small amount of stock to run that and check if the system was able to output the data we wanted. I recalled we started with only six stocks in our system. After compiling social trends on these stocks, we put them into the algorithms we used to form the graph and compute the necessary computations. Now, the graph would display the correct numbers regarding the price and social trends of the stock. Training machine learning also took large amounts of data. We would have to determine if the output data matches the current trend and is likely going to display an accurate prediction. After analyzing the graphs and the data, we were able to conclude that the prediction modeled highly accurate predictions. With multiple trials implemented, we were able to begin implementing more stocks into the system for computation and data extraction. Once we had one algorithm working successfully, the successive runs would almost always predict accurately.

The rest of the paper will follow a similar format to most research papers. Section 2 explains the challenges we faced when building our app and testing for the experiment; Section 3 dives deep into the details of the program, including codes and solutions we came up with to resolve the challenges identified in part 2. We will be showing the structure, and the diagram of meaningful tests methods we have used; Section 4 contains the specific experiment we have done, including an evaluation and analysis of the results; Section 5 lists many relevant works that have already been done by someone else regarding stock investing and stock forecasting. This section compares our work to others' work; finally, Section 6 will conclude our study and state any future plans and improvements we may continue to target.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Implementing Large Numbers of Stocks

In order for our app to be as professional as possible, we need to store almost all the stock in the stock market in our database. Unlike those large trading apps, we do not have a huge cloud to store all these stocks compatibly. We had to write the tickers for these stocks manually. We tried some of them with our current system ability, but it turned out that on the user's end, it took forever to load, and it caused a huge lag for the user. Also, if we would successfully implement these 6000 stocks into our backend, it would be extremely hard to display to the users in the frontend. This part of the process took us the longest time in our build, but we were able to minimize the lag by only displaying a small number of stocks in the search bar and displaying the rest when users intentionally search them up using the search bar. We noticed a huge decrease in lag and more fluency on the screen.

2.2. Creating a Convenient Frontend

We were mainly concerned with the user's experience when developing this app. The method we used when creating the user frontend was to use Android Studio to program the screen using Dart language. We tried to implement many buttons and clickable features, but some of them turned out to be unresponsive. We found some people to test out the app, and many of them reported that some of the buttons do not work and functions such as watchlist and daily trend do not work. There was also a problem with the back button at some point, which in turn caused users to have to reopen the app to be reset to the home screen. It was mainly due to the problem with our code, so we had to go back to parts of the code and locate where the error is occurring. On top of fixing it, we also had to continuously brainstorm ways to make the app more inclusive. In fact, we are still in the process of improving our UE by adding other useful features.

2.3. Stock Performance Prediction

The biggest selling point of our app would probably be the ability to predict the performance of a state and output a percent increase or decrease. This inevitably becomes the biggest challenge in our product. Our program consists of AI machine learning using a modeled representation as a polynomial function graph. We are performing our prediction based on the number of posts on social media about a specific stock. As you might know, participation on social media is very unpredictable every day. Even if we use machine learning to study the day the previous day, it will still be hard to accurately present the data the next day without a huge margin of error. This is something that we would have to dive deeper into in terms of long-term improvement. We need to research and choose the closest related modeling technique for the machine and hope the get the closest result to the actual trend.

3. Solution

This app is designed to provide users with a percentage increase or decrease of a specific stock price in the next three days based on the data retrieved from social media. There are three major components that formulate the entire app: the user frontend, the backend database, and the machine learning algorithm. First off, the entire front end was created using Dart in Flutter. Using this software allows us to develop an app accessible to both IOS and Android users [7]. Features include displaying the compiled data to the user in the simplest manner and being able to adapt to changes made in the backend quickly. Another software we used during the build was Google Firebase. Google firebase does a magnificent job at storing large amounts of data and providing analysis of these data [8]. We were able to store data inside the firebase that could later be used to connect to our backend to be computed. Along with Google Firebase, we also used Amazon AWS, which is a cloud platform that provides effective cloud computing solutions [9]. Since we have more than 6000 stocks in our database, we needed an efficient cloud computing software that could deliver the analysis of all the 6000+ stocks we need. Last but not least, the entire backend and machine learning functionality are created using Python. The Python server manages all the data that will be inputted into the app to make predictions about the growth. We have implemented both the linear regression model and polynomial model to represent the prediction that will be displayed correctly.

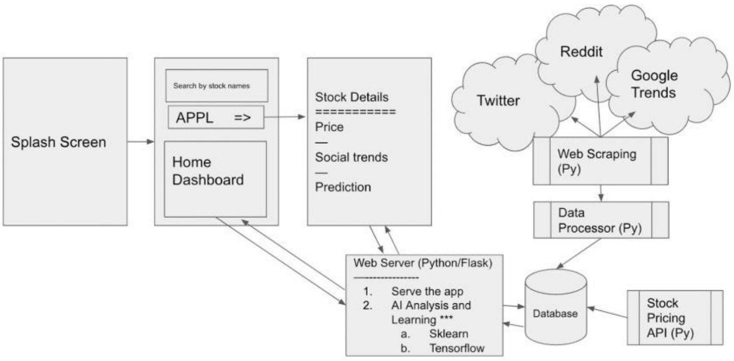


Figure 1. Overview of the solution

1) Web Scraping

    a. Twitter API code + text description
    b. Yahoo finance API code + text description

2) Server/Database

    a. Database
        i. Screenshot
        ii. text to talk about how it works;
        iii. the data structures;

    b. Server
        i. Python Flask code + text description
        ii. AWS and deployment (screenshot) + text description

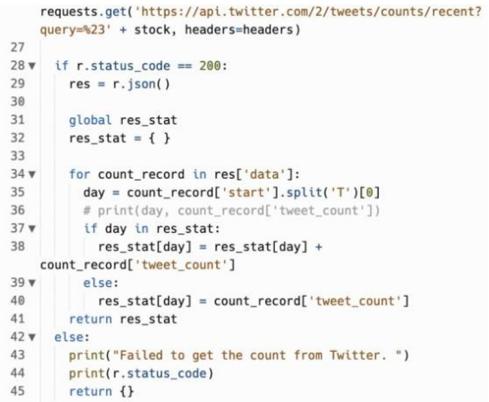
3) AI/Machine Learning

    a. scikit-learn library
    b. linear regression etc.

4) App

    a. screenshots
    b. find 1-2 important code excerpts
    c. text to talk about how the app works and what the code do

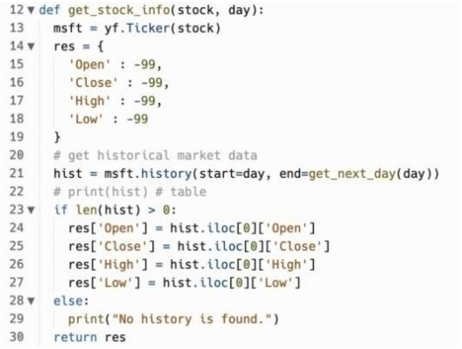
The back end of the app contains the web scraping algorithm, servers/databases, and machine learning models. In order to make an accurate prediction, the machine learning server and database have to be constantly communicating with each other. The AI learns from the data provided by the cloud to output results that predict the future trend. These data all came from web scraping Twitter and Yahoo Finance.



```
requests.get('https://api.twitter.com/2/tweets/counts/recent?
query=%23' + stock, headers=headers)

  if r.status_code == 200:
    res = r.json()

    global res_stat
    res_stat = { }

    for count_record in res['data']:
      day = count_record['start'].split('T')[0]
      # print(day, count_record['tweet_count'])
      if day in res_stat:
        res_stat[day] = res_stat[day] +
count_record['tweet_count']
      else:
        res_stat[day] = count_record['tweet_count']
    return res_stat
  else:
    print("Failed to get the count from Twitter. ")
    print(r.status_code)
    return {}
```

Figure 2. Twitter API application code

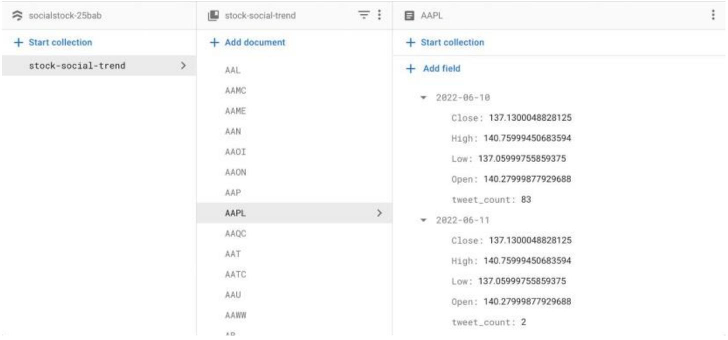
One of the major components of the design is web scarping. This figure shows the utilization of Twitter API to scrape all the data and tweets and Twitter on stocks. These codes are written so that we are able to examine the tweet count from different days of the week. If the system failed to scrape the data from Twitter, then a failure message will be displayed so we can fix it accordingly. The reason why this is an essential part of the program is that we need this data to formulate our prediction. Without this data, there is no way we could move forward with our app. The app is centered around social media trends, so scraping data from social media platforms is necessary. Later, this information will connect with our machine learning model to be analyzed even further.



```
def get_stock_info(stock, day):
  msft = yf.Ticker(stock)
  res = {
    'Open' : -99,
    'Close' : -99,
    'High' : -99,
    'Low' : -99
  }
  # get historical market data
  hist = msft.history(start=day, end=get_next_day(day))
  # print(hist) # table
  if len(hist) > 0:
    res['Open'] = hist.iloc[0]['Open']
    res['Close'] = hist.iloc[0]['Close']
    res['High'] = hist.iloc[0]['High']
    res['Low'] = hist.iloc[0]['Low']
  else:
    print("No history is found.")
  return res
```

Figure 3. Yahoo Finance API application code

This figure is the other part of web scraping included in our program. This part is relatively straightforward. We scarped stock information, including the opening and closing price of the stock, as well as the highest and lowest price of the stock during the day. Again, we used Yahoo Finance API to get the information of the stock and scrape relevant information granted with the API code. If no history is found, an error message will be displayed. This part of the algorithm is also crucial in our app because the basic information about the stocks will be presented to the users when they choose one to look at. This information will be stored in Google Firebase and connected straight to the app's front end.



stock-social-trend

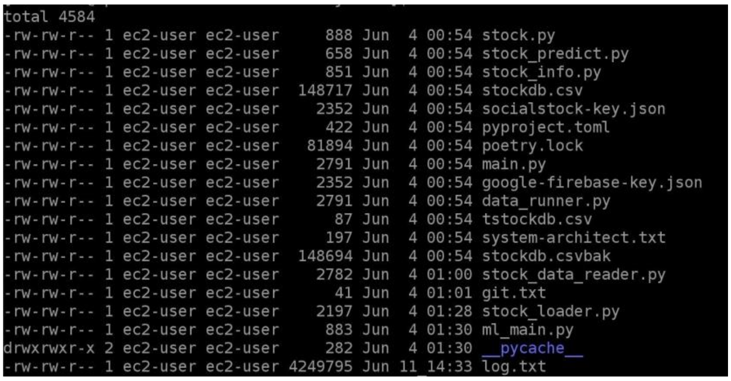
AAL
AAMC
AAME
AAN
AAOI
AAON
AAP
AAPL
AAQC
AAT
AATC
AAU
AAWW

▾ 2022-06-10
Close: 137.1300048828125
High: 140.75999450683594
Low: 137.05999755859375
Open: 140.27999877929688
tweet_count: 83

▾ 2022-06-11
Close: 137.1300048828125
High: 140.75999450683594
Low: 137.05999755859375
Open: 140.27999877929688
tweet_count: 2

Figure 4. Google Firebase data storage

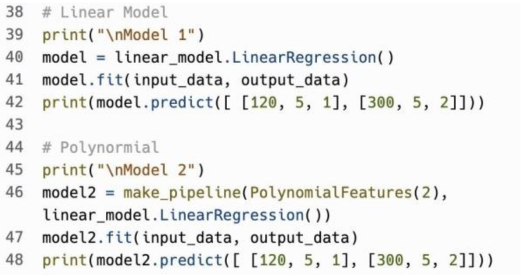
We used Google Firebase to create bigger storage to store all the trends scaped from social media and financing websites. The Firebase contains a list of all the stocks we will have in our app. Inside those stocks, we have a list of information saved regarding these stocks. We pulled all this information from Yahoo Finance which is a reliable source that displays the basic information about the stocks. Figure 4 is just a screenshot taken from the database in Google Firebase. Information includes opening and closing prices, the highest and lowest prices, and the date when this information was retrieved.



```
total 4584
-rw-rw-r-- 1 ec2-user ec2-user     888 Jun  4 00:54 stock.py
-rw-rw-r-- 1 ec2-user ec2-user     658 Jun  4 00:54 stock_predict.py
-rw-rw-r-- 1 ec2-user ec2-user     851 Jun  4 00:54 stock_info.py
-rw-rw-r-- 1 ec2-user ec2-user  148717 Jun  4 00:54 stockdb.csv
-rw-rw-r-- 1 ec2-user ec2-user    2352 Jun  4 00:54 socialstock-key.json
-rw-rw-r-- 1 ec2-user ec2-user     422 Jun  4 00:54 pyproject.toml
-rw-rw-r-- 1 ec2-user ec2-user   81894 Jun  4 00:54 poetry.lock
-rw-rw-r-- 1 ec2-user ec2-user    2791 Jun  4 00:54 main.py
-rw-rw-r-- 1 ec2-user ec2-user    2352 Jun  4 00:54 google-firebase-key.json
-rw-rw-r-- 1 ec2-user ec2-user    2791 Jun  4 00:54 data_runner.py
-rw-rw-r-- 1 ec2-user ec2-user      87 Jun  4 00:54 tstockdb.csv
-rw-rw-r-- 1 ec2-user ec2-user     197 Jun  4 00:54 system-architect.txt
-rw-rw-r-- 1 ec2-user ec2-user  148694 Jun  4 00:54 stockdb.csvbak
-rw-rw-r-- 1 ec2-user ec2-user    2782 Jun  4 01:00 stock_data_reader.py
-rw-rw-r-- 1 ec2-user ec2-user      41 Jun  4 01:01 git.txt
-rw-rw-r-- 1 ec2-user ec2-user    2197 Jun  4 01:28 stock_loader.py
-rw-rw-r-- 1 ec2-user ec2-user     883 Jun  4 01:30 ml_main.py
drwxrwxr-x 2 ec2-user ec2-user     282 Jun  4 01:30 __pycache__
-rw-rw-r-- 1 ec2-user ec2-user 4249795 Jun 11 14:33 log.txt
```

Figure 5. Amazon AWS cloud computing process

On top of the Google Firebase we used for storage, we also used Amazon AWS for cloud computing and analysis. In order to smoothly implement the 6000+ stocks into our app, the information needs to be computed constantly at the backend. We chose not to put all 6000+ stocks directly available to the users in the front to avoid lags, but they do exist and are consistently being computed with AWS. With AWS cloud computing, more stages and analyses of the data can be delivered and be available to the users upon request. Figure 5 is a screenshot of the server taken at a random time, but it shows the data being computed continuously.



```
# Linear Model
print("\nModel 1")
model = linear_model.LinearRegression()
model.fit(input_data, output_data)
print(model.predict([ [120, 5, 1], [300, 5, 2]]))

# Polynormial
print("\nModel 2")
model2 = make_pipeline(PolynomialFeatures(2),
linear_model.LinearRegression())
model2.fit(input_data, output_data)
print(model2.predict([ [120, 5, 1], [300, 5, 2]]))
```

Figure 6. Machine learning model code

Here comes the machine learning part of the program, an essential area of focus that outputs the final prediction. Figure 6 shows the code we have written while programming machine learning. We used the scikit-learn library to construct two models that we will be using when making predictions. Both of these models require construction for a line of best fit. For the linear regression model, some values are inputted for the machine to construct a line of best fit in a linear manner. For the polynomial regression model, other values are inputted to construct a line of best fit in a curvilinear manner. Once the structure of the model is formed, the machine will learn from its previous data to output a prediction for data not part of the model.

Lastly, the front end of our app was constructed with Flutter to display all the functionalities to the users in a friendly and convenient manner.

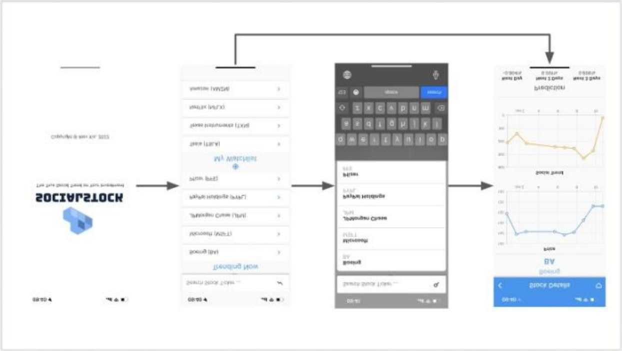
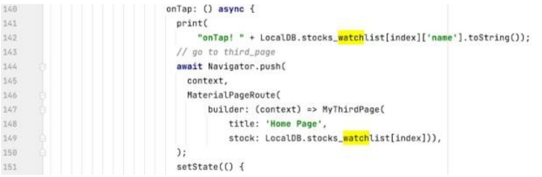


Figure 7. App storyboard

Figure 7 is a diagram of the front end of the app. When prospective users download the app, their interfaces will look like the one shown here. There will be a splash screen with the logo and name of the app, along with the copyright. The splash screen disappears after two seconds, and the user will be prompted to enter the home page of the app. On the home page, there are three components. The first one is a search bar where users can search all the 6000+ stocks we have in our backend. They can simply type the ticker or the name of the company and select their stock. The second component is a daily trend of the stocks that are most discussed on social media recently. The third component is the ability to save a stock to the user's watchlist for easier access. Finally, once the user decides on a stock that they want information on, a detailed information page will be displayed that contains the stock name, price, trend, and, most importantly, its predictions.



```
onTap: () async {
  print(
      "onTap! " + LocalDB.stocks_watchlist[index]['name'].toString());
  // go to third_page
  await Navigator.push(
    context,
    MaterialPageRoute(
        builder: (context) => MyThirdPage(
            title: 'Home Page',
            stock: LocalDB.stocks_watchlist[index])),
  );
  setState(() {
```

Figure 8. Watchlist function code with flutter

Figure 8 is just a snapshot of one of the functions we included in our app, which is a watchlist [15]. Many online investing apps have this function, so we included it to improve the user experience. How this works is that each user has their own watchlist that they can manage. Whenever they search for stock or click on a stock, the user has the ability to add that stock to their watchlist by clicking a button on the top right corner. There is nothing special about the function, just to make the app more inclusive.

4. EXPERIMENT

4.1.Experiment 1: The Accuracy siung Different Machine Learning Algorithms

For all the three experiments we have conducted, we used a total of 50 input data that contains information on ten stocks we picked. Each stock contains five lines of information with four columns: trend count, month, day, stock No. This input data remains the same for the first two experiments. For the last experiment, we have to cut down the number of columns in order to meet the criteria of the experiment.

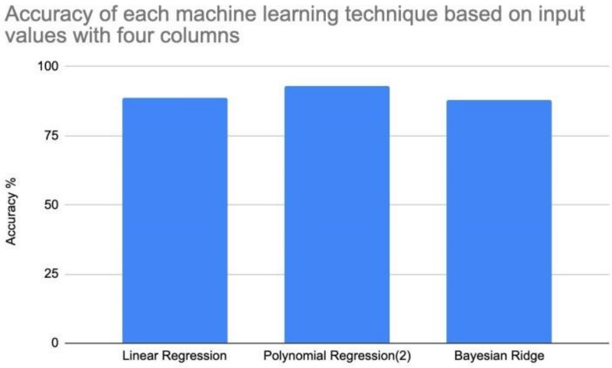


Figure 9. Graph displaying the accuracy of different machine learning algorithms

Clients may constantly be suspicious of the actual precision of the app. To address that, we have to test out the reliability of our machine learning program. The first step in creating a machine learning program is to select which machine learning method is most appropriate to use to compute our data. There are many learning techniques available, but only some are designed to solve our circumstances. After researching, we have picked three machine learning techniques that would yield the highest accuracy. This first experiment we performed was to simply test out which of the selected machine learning techniques would be most accurate when it comes to predicting great numbers of inputs. All three models showed slight variation, with polynomial regression being the most accurate. The concept of polynomial expression is relatively broad. Different degrees yield different results. In this experiment, we only used polynomials to the second degree for the simplest. We will be testing polynomials to different degrees in our second experiment. One reason for this outcome could be because the data we used neither shows a linear pattern (linear regression) nor shows a huge fluctuation (bayesian ridge). The result of this experiment informs us that the parabolic shape generated by the polynomial regression to the second degree most appropriately represents the trend that the stock information contains. This is not to say that the other two methods are not accurate; they are still pretty accurate based on the result but might only work in less likely conditions which stock data we are analyzing.

4.2. Experiment 2: The Accuracy of Using Different Parameters for Polynomial Regressions

To know if our user experience is good and if the user is satisfied with the UI design, we publish our app in the market and advertise the production in communities and schools. A total of more than 1,000 students used our app to help their parents, and we received a total of 20 feedback questionnaires. We collect the data and make a diagram to show the feedback result.

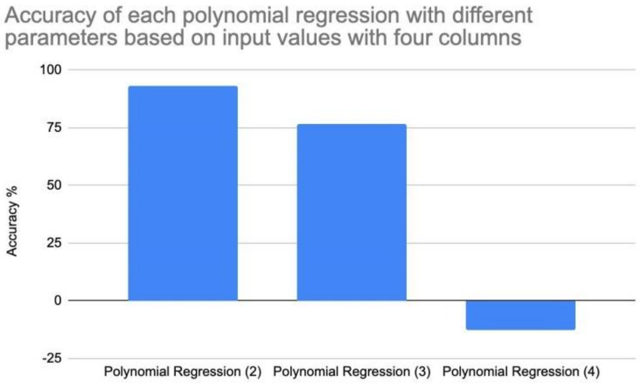


Figure 10. Graph displaying the accuracy of each polynomial regression with different parameters

Since we tested that polynomial regression yields the best result out of all the three methods, we went even further to test out polynomial regression with different parameters. The word polynomial means an equation with several terms and different powers. In order to make even more accurate results to improve the app, testing out which degrees predict the best result is necessary. In this experiment, we tested polynomials with degrees of 2, 3, and 4. As a brief overview of what a polynomial graph looks like in math, any polynomial with an even degree resembles the shape of a parabola, and any polynomial with an odd degree resembles the shape of a snake. Given the two completely different shapes a polynomial graph represents, the outcome of the result would be a lot different. As shown in the chart above, we tested out that a polynomial regression with a degree of 2 predicts our given input the best. Polynomial regression with a degree of 3 did show some consistency when predicting the outcome, so it is still a reliable algorithm but might not be as inclusive as the former choice. What is interesting about this experiment is the fact that a polynomial regression with a degree of 4 showed negative results when the score came out. Theoretically, a polynomial with a degree of 4 looks very similar to a polynomial with a degree of 2. The only difference between them is that a degree of 4 makes the parabola sharper. In other words, it makes the parabola closer to a U shape with noticeable vertices. A reason behind this might be that the sharp turn created by the high degree limits the condition of the data. Since the data almost never follows the domain of the graph, stricter conditions underperform.

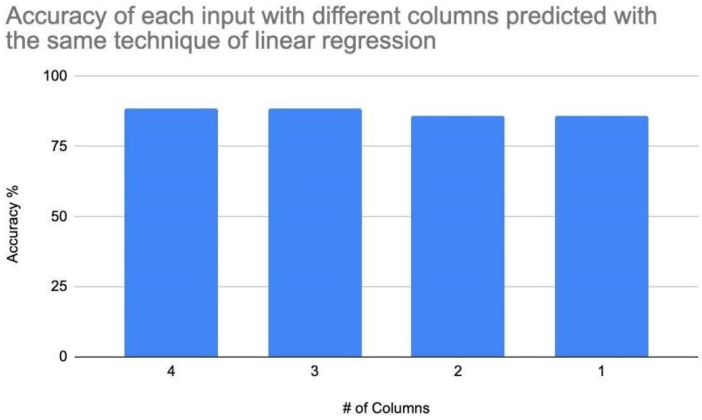


Figure 11. Graph displaying the accuracy of datasets with different columns

Last but not least, testing the suitable number of columns in the data finalizes our machine learning technique. When we web scraped the information on a certain stock, we stored many data such as trend count, price, date, etc. in our storage. One set of data contains many indexes. In terms of testing our algorithm, we picked at most four values to include that are the most relevant to the stocks. Those four values we picked are trend count, month, day, and the serial number we gave for each stock. When each of these values is included in a stored list, the machine will examine each index and study them with the pre-given performance to come up with a result. This means that when one index is removed, the result of the prediction will change. The order given before is listed from most important to least important (trend count, month, day, serial number). With every trial, we removed the index from the right side until we reached only one index of the stock's trend count. Our initial hypothesis was that with fewer indexes, the algorithm would perform more accurately since it does not have to focus on a huge dataset. However, from the final result, it is clear that this hypothesis is turned down. Instead, there is a clear pattern that when the algorithm is given more information, it ends up performing better, even only a slight bit.

5. Related Work

Ding, Xiao, et al. proposed a different technique when predicting stock performance [10]. This was an event-driven technique in which developers used many past events as an indicator of how the stock market is likely going to perform in the future. This program extracts news events from outside sources that could be summarized into an inclusive prediction. Our work differs from this individual's since we were extracting social trends such as tweets and posts about a certain stock. By focusing on social trends instead of solely on news events, the data can be even more inclusive because social trends are directly linked to news and many other factors. The social trend represents investors' expectations toward a stock which came from influences like news events.

Holthausen, Robert W., and David F. Larcker used an old-fashion way of predicting stocks [11]. They only looked at the prices of the stocks from many years ago as their evidence for prediction. This method is done by deeply analyzing the price changes, including many performance indicators that give the developer the most appropriate outcome. This is a very straightforward method and probably the simplest, but it has its flaws. As most people are aware, the stock market is full of unpredictable factors. Looking at previous performance only serves as a reference; it does not accurately predict its performance in the future. In our program, we analyze data from a wide spectrum and consider many different trends before reaching a conclusion. The social trend has more reliable information since those analyses essentially came from professionals that knew the topic well. We pull those data to help us formulate our program.

Last but not least, Rather, Akhter Mohiuddin, Arun Agarwal, and V. N. Sastry used a hybrid model of representation for machine learning and big data analysis [12]. In this design, the developer used multiple machine learning techniques far beyond linear and polynomial regression. Many non-linear models were used and tested to improve the accuracy of the predicted result. When designing our work, we omitted using multiple techniques because it requires endless hours of testing and reformulating to match a technique to a stock. We simply tested and found the most effective method that covers most of the trends that came from those 6000+ stocks in our database. This not only made the process more efficient but also alleviated the pressure placed on the algorithm.

6. CONCLUSIONS

As technology advances and people invest more, virtual currency and digital exchange will be the trend in almost every nation [13]. However, professional investors oftentimes dominate the market, suppressing amateur investors and rug-pulling on average citizens [14]. In order to create a fair economy, it is important that everybody uses their wealth to the fullest potential. By providing an accurate prediction of the performance in the stock market, we can help amateur investors to quickly and safely adapt to the complicated market and achieve multiplication of wealth. We web scraped data from Twitter to use as our concrete evidence for the prediction. Social media trends directly reflect on the future performance of a stock because, with a high number of discussions on a certain stock, people's expectations rise, thus influencing the performance of the stock. By using the machine learning technique to analyze these social trends, the system is able to predict as close to precision as possible. With this reliable source as a backbone supporting the prediction, users can feel free to use our product without any major concerns. Additionally, our implementation of all the 6000+ stocks creates many more choices for the users. We try to aim for a perfect user experience; providing the basic quantity of stocks is the first step in servicing our users. Our databases safely store these stocks, and all the information on these stocks came from Yahoo Finance. Users can almost find any public stocks. Together with our easy-to-use front end, people from all age groups will feel effortless when using the app.

Even though we have tested our machine learning program thousands of times, accuracy is still a major concern for us. It is reasonable to assume that when it comes to dealing with prediction, there is not a single way to predict the future 100% accurately. Although our model represents the performance of a stock very closely, the margin of error is still present and sometimes bigger on certain stocks. Second, renewing the stock list periodically will be hard to achieve. As more private companies become public and enter the stock market, more stocks will appear on exchange platforms. Our program currently does automatically renew the stock list, it must be done manually by the developers.

Along with updating the app from time to time, it is very important for us to keep improving the accuracy of the prediction. Having a slightly inaccurate prediction can hurt our users. The core concept of this app is to predict. If we cannot effectively accomplish that, the rest of the functionalities are just extras.

REFERENCES

[1] Stockhammer, Engelbert, and Lucas Grafl. "Financial uncertainty and business investment." Review of Political Economy 22.4 (2010): 551-568.

[2] Ramazani, Jalal, and George Jergeas. "Project managers and the journey from good to great: The benefits of investment in project management training and education." International Journal of Project Management 33.1 (2015): 41-52.

[3] Liu, Jun, and Francis A. Longstaff. "Losing money on arbitrage: Optimal dynamic portfolio choice in markets with arbitrage opportunities." Review of Financial studies (2004): 611-641.

[4] Groth, Sven S. "Does algorithmic trading increase volatility? Empirical evidence from the fully-electronic trading platform Xetra." (2011).

[5] Kieffer, Christine, and Gary Mottola. "Understanding and combating investment fraud." Financial decision making and retirement security in an aging world 185 (2017).

[6] Mazorra, Bruno, Victor Adan, and Vanesa Daza. "Do Not Rug on Me: Leveraging Machine Learning Techniques for Automated Scam Detection." Mathematics 10.6 (2022): 949.

[7] Watanabe, Y., et al. "An experimental study of paper flutter." Journal of fluids and Structures 16.4 (2002): 529-542.

[8] Chatterjee, Nilanjan, et al. "Real-time communication application based on android using Google firebase." Int. J. Adv. Res. Comput. Sci. Manag. Stud 6.4 (2018).

[9] Bermudez, Ignacio, et al. "Exploring the cloud from passive measurements: The Amazon AWS case." 2013 Proceedings IEEE INFOCOM. IEEE, 2013.

[10] Ding, Xiao, et al. "Deep learning for event-driven stock prediction." Twenty-fourth international joint conference on artificial intelligence. 2015.

[11] Holthausen, Robert W., and David F. Larcker. "The prediction of stock returns using financial statement information." Journal of accounting and economics 15.2-3 (1992): 373-411.

[12] Rather, Akhter Mohiuddin, Arun Agarwal, and V. N. Sastry. "Recurrent neural network and a hybrid model for prediction of stock returns." Expert Systems with Applications 42.6 (2015): 3234-3241.

[13] Dibrova, Alina. "Virtual currency: new step in monetary development." Procedia-Social and Behavioral Sciences 229 (2016): 42-49.

[14] Gaskill, N. "Rorty against Rorty: climate change, rug-pulling, and the rhetoric of philosophy." Common Knowledge (2020).

[15] Bannier, Christina E., and Christian W. Hirsch. "The economic function of credit rating agencies–What does the watchlist tell us?." Journal of Banking & Finance 34.12 (2010): 3037-3049.

AN INTELLIGENT MONITORING AND REWARDING SYSTEM TO ASSIST THE TENNIS TRAINING USING SENSOR-BASED DETECTION AND IOT SYSTEM

Andy Kuang[1] and Yu Sun[2]

[1]Eleanor Roosevelt High School, 7447 Scholar Way, Eastvale, CA 92880
[2]California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

ABSTRACT

Coronavirus started in 2019 and it's still a major problem today [1]. This disease led to the start of pandemics around the world, in which some students are still using online learning platforms today, and their guardians leaving them unattended to provide for their families [2]. However, with the lack of supervision children are taking advantage of these times to perform unproductive activities such as gaming. During school days, there are also many breaks provided for students to relax and reset their mentality, which allows a student to be focused during class, but this doesn't seem to be the case students are spending this time indoors after many hours of staring at a device, instead spending it outdoor can relax their eyes also preventing eye damage. This paper proposes software that tracks a student's productivity based on their tennis racket movement and speed using a particle board, accelerometer, and tracker [3]. With a tracker, guardians would be able to get constant updates on their children's activities. We applied our application to a real-life scenario and conducted a qualitative evaluation of the approach. The results show that students spend less time indoors performing nonproductive activities, students spend more time outside playing their sport of desire, and parents are less stressed about their children's educational and physical health.

KEYWORDS

Pandemic, C++ and HTML, Mobile APP.

1. INTRODUCTION

The education of a student is very important, it is a key factor that makes a person successful in the future. However, a deadly disease known as the Coronavirus, sent our world into a global pandemic, and many schools diverted to online teaching, but there are many wrongs about online teaching [4]. For example, due to the long amount of screen time, eye damage could be a possibility [5]. Even though schools provide many breaks and passing periods for students to relax and stop looking at their screens, many students don't take this seriously. Instead of spending this time outside taking a breather or playing a sport, students are spending this time indoors playing video games or going on social media. Another factor contributing is the number of distractions, because students are at home, teachers really don't have an idea of what they are doing. Study shows that about sixty percent of the class don't pay attention to the class, instead spending this time doing other things. By solving this problem parents and teachers would no

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 177-184, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121614

longer have to worry about their students and children facing potential physical or mental health issues. Parents also no longer have to worry about their children not learning the curriculum.

Some of the related systems and tools that have been proposed to the public are security cameras, which allow the users to get live updates on their houses [6]. Another tool that has been proposed to the public is fitness tracker watches, which provide the user with a daily report of their activities, step count, heart rate, etc. However, these proposals assume that their users are productive and use their product as intended, which is rarely the case in practice. Furthermore, there are many ways to beat the system. For example, security cameras can be moved to a different area where they lose visuals of the intended section of the house, and fitness tracker watches could be taken off and can no longer track your daily activities [15]. Their implementations are also limited in scale. Some studies had also found methods in keeping student production. For example, some studies found that in order to keep yourself from doing unproductive activities you must obtain a study habit, but this rarely seems to be the case, students/children are often distracted because of something they want or something they want to do.

In this paper, we follow the same line of research by demonstrating our product to answer the previous problems mentioned. Which is, How can we reduce the number of unproductive activities being performed by students and prevent them from being distracted during their online learning sessions, while also balancing their indoor activities and their outdoor hobbies [7]. Our goal is for our product to show positive results in the education market, and also help aid those families that are struggling during these times. Our method is inspired by the discussions of parents during this pandemic, many of them mentioning how their children are being distracted during online learning, which results in them failing classes. There are many good features of our product compared to existing tools available. First, compared to existing tools available to the public, our product will be able to track the user's activity by tracking the speed at which the racket is moved. Second, our product will have a guardian alarm, whenever the user is doing something unproductive a message would be sent to their guardian. Third, their daily performance would then be displayed on their own profile when signed up with our website, and their activities would be analyzed providing a better outcome. Therefore we believe that our product will be the answer to the number of nonproductive activities being performed during this pandemic.

Our product went through various scenarios of testing in order to receive positive results, we demonstrated how the above combination of techniques decreases the number of unproductive activities being performed at home and balanced it with each child's favorite outdoor activity. First, we show the usefulness of our product by testing it among various families that felt like their children were falling behind in their classes due to the number of distractions and the unproductive activities being performed. The student's activity would be tracked whenever the designated sport they want to play is in action. For example, if a student wants to play tennis, then our product would be attached to it and the speed at which the racket is moved would be tracked and logged to the firebase [8]. The activity would then be sent to their guardians. Our product could also be attached to different sport-related items, such as a baseball bat, golf club, etc. After a majority of families tested our product, the results show that students are able to balance their day between indoor activities and outdoor activities, therefore reducing the amount of device screen time and also reducing the possibility of potential eye damage. Second, we also analyzed the results being sent to each guardian, the results show that their children spend a few weeks with the product, and spending time outdoors, spending time outdoors starts to become a habit, and also show more engagement in their classes.

The rest of the paper is organized as follows: Section 2 describes the details of the challenges that we faced during the experiment and the design of our product; section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in section 2; section 4 presents the related details on the experiment we did, following by presenting the related work in section 5. Finally, section 6 will conclude our study and experimentation as well as point out the future works on our project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Keeping our Prototype Attached to the Item

One major problem we had during our testing phase was attaching our prototype to an item without it getting knocked off, damaged, or affecting the player. For example, if we had a tennis racket we tried attaching it above the handle where it could calculate the speed of the racket when swung more effectively. However, attaching it to the racket could lead to multiple problems. First, if we decided to attach it to the rim of the tennis racket there's a high possibility our product will fly off when swung, and distract the player. Another problem if we placed our product on the rim, it could resist the air leading to a slower tennis swing. After multiple tests on the placement, which negates all of the above problems the best, was attaching our prototype to the handle. When our product was attached to the handle of the racket, it seemed to prevent air resistance, the chances of it falling off were slim, tracking the speed was still sufficient, and it felt less distracting.

2.2. Sending the Data to the Firebase

Another major problem we dealt with during the production of our prototype was setting up the firebase. Creating the firebase was one thing, but constantly sending the data to our website is another. We had to figure out multiple code strings in order for our data to be sent to the firebase and work properly. However, by sending the data constantly to the firebase we started receiving restrictions, at one point when the accelerometer was in action, not a single data point was being recorded. To fix this we started lowering the amount of time the data was being tracked from once every second to once every minute. Another problem we had with the firebase was the connection and power. In order for our accelerometer to track the data and send it toward our firebase, it needed a stable connection and battery. Instead of needing a wire to charge and send a connection to our database, we decided to use a battery attachment. This battery provided both constant power and connection.

2.3. Beating the System

Our product basically reduces the amount of nonproductive activity being performed at home by placing a tracker on a sports item/tool and urging children to spend time outside more by balancing the time and activities being spent indoors and outdoors. However, many children figured out a way to beat the system, instead of actually playing the sport, they would take the item and just swing it around every minute to keep the tracker going, therefore the data being tracked would be sent to their guardian's email, leading to the guardians believing they are actually spending time outside when in reality they are just indoors playing games or going on social media. In order to fix this problem, the code string tracking the speed for when a specific speed is reached is raised higher. Therefore, children can't just swing the racket at ten mph and the data would be tracked instead they would have to perform regular tennis swings around 80 mph.

3. Solution

Our product (Smart Racket) is a software and application that allows a guardian to track their children's activity. Smart Racket could be represented as a nanny cam but has a greater effect on children. During online school learning, students are given many breaks and many students decide to play video games or go on social media during this time even though they may have just attended five-hour learning sessions, without these breaks from their digital screens may cause eye damage or even worse blindness, because of this our product was designed [9]. During these breaks, if students decide to continue with their nonproductive activity, emails would be constantly sent to their guardians. In order to track the student's outdoor activity, the Boron and Accelerometer would be attached to a student's tennis racket. Whenever the tennis racket is in motion the accelerometer would calculate the speed at which the racket is swung. When the racket reaches that speed the data would then be sent to the firebase, then transferred to our website. Our application is written using C++ and HTML. There are four main components of our product. First, the particle networking development tool (Boron) is the main way to keep our project running in areas where there is no WIFI or unreliable internet. Second, the Accelerometer tracks the speed and movement of the tennis racket when swung and sends the data to the firebase. Third, is the firebase, which stores and logs all the data collected from the Accelerometer. Lastly, the Website, there are multiple web pages inside our website, a profile tab, home page, data tab, and much more [10]. The visual below describes these components in detail and also gives a great visualization of how our product works.

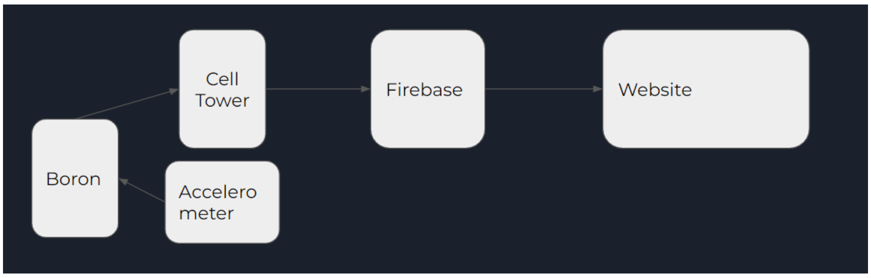
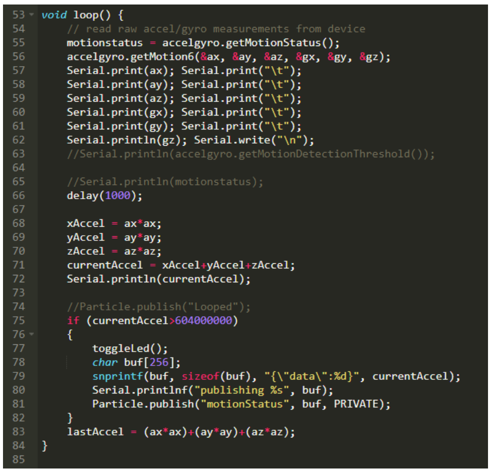


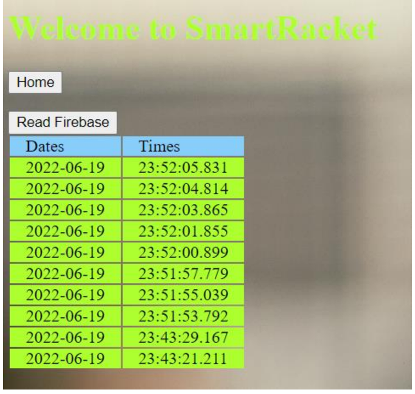
Figure 1. Overview of the solution

During the production of our product different tools, algorithms, and services were used to create our final product. Tools such as a Particle Boron, Accelerometer, Tennis Racket, etc. Services such as Replit, Particle, etc. Our coding is written in C++ and HTML. Image A shows part of our coding implemented into the accelerometer in order to track an average tennis racket swing. However, in order to start tracking the speed at which the racket is swung, we had to start from the basics, by calculating gravity and the minimum speed at a tennis racket is swung normally which is around 60 mph. When the racket is in motion and the X-axis Acceleration + Y-axis Acceleration + Z-axis Acceleration doesn’t reach the minimum speed that the racket is supposed to swing at then the data wouldn’t be logged. However, if it does exceed the limit then the current Acceleration at which the racket is swung would be logged. In image 2, you can see part of our website design for how the data looks every time a data point is logged. However, the website is still in production and it isn’t fully developed yet. Currently, we are implementing different ideas into our website such as a profile page, login credentials, etc.



```
void loop() {
    // read raw accel/gyro measurements from device
    motionstatus = accelgyro.getMotionStatus();
    accelgyro.getMotion6(&ax, &ay, &az, &gx, &gy, &gz);
    Serial.print(ax); Serial.print("\t");
    Serial.print(ay); Serial.print("\t");
    Serial.print(az); Serial.print("\t");
    Serial.print(gx); Serial.print("\t");
    Serial.print(gy); Serial.print("\t");
    Serial.println(gz); Serial.write("\n");
    //Serial.println(accelgyro.getMotionDetectionThreshold());

    //Serial.println(motionstatus);
    delay(1000);

    xAccel = ax*ax;
    yAccel = ay*ay;
    zAccel = az*az;
    currentAccel = xAccel+yAccel+zAccel;
    Serial.println(currentAccel);

    //Particle.publish("Looped");
    if (currentAccel>604000000)
    {
        toggleLed();
        char buf[256];
        snprintf(buf, sizeof(buf), "{\"data\":%d}", currentAccel);
        Serial.printlnf("publishing %s", buf);
        Particle.publish("motionStatus", buf, PRIVATE);
    }
    lastAccel = (ax*ax)+(ay*ay)+(az*az);
}
```

Figure 2. Screenshot of code



Welcome to SmartRacket

Dates	Times
2022-06-19	23:52:05.831
2022-06-19	23:52:04.814
2022-06-19	23:52:03.865
2022-06-19	23:52:01.855
2022-06-19	23:52:00.899
2022-06-19	23:51:57.779
2022-06-19	23:51:55.039
2022-06-19	23:51:53.792
2022-06-19	23:43:29.167
2022-06-19	23:43:21.211

Figure 3. Screenshot of homepage

4. EXPERIMENT

4.1. Experiment 1

In order to accurately show our results, we collected 300 data sets from 10 different guardians that have experienced this problem with their children. Most of these children are spending their time during this pandemic inefficiently. In order to show the results of our product, we conducted multiple experiments to test and verify our product: How long a student spent outside per two days out of 10 different children, the number of students that actually spent time outside compared to the number of students that stayed indoors. After all, the experiments were completed, the guardians of each child/student were satisfied with our product, most of them saying their children were spending time outside learning how to play tennis with their friends.

A. The Number of Students That Went Outside vs Day #

Figure 4 illustrates the number of times a student went outside per day regarding the weekends. As you can see in the graph, in the beginning, when each student was given our product, not that many students took it seriously. However, as time passed, positive results started showing, students were going outside more as days passed. Showing the effectiveness of our product.

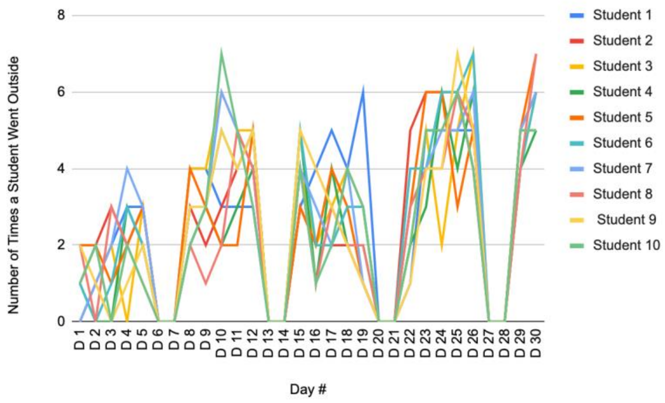


Figure 4. The Number of Students That Went Outside vs Day #

4.2. Experiment 2

Figure 5 shows how long each student went out in minutes for every two days. As the same as figure 4, in the beginning, students and children didn’t spend much time outside, most of them around 30 minutes every 2 days. However, as time passed students started to spend around 80 minutes outside per 2 days. Showing the effectiveness of our product.

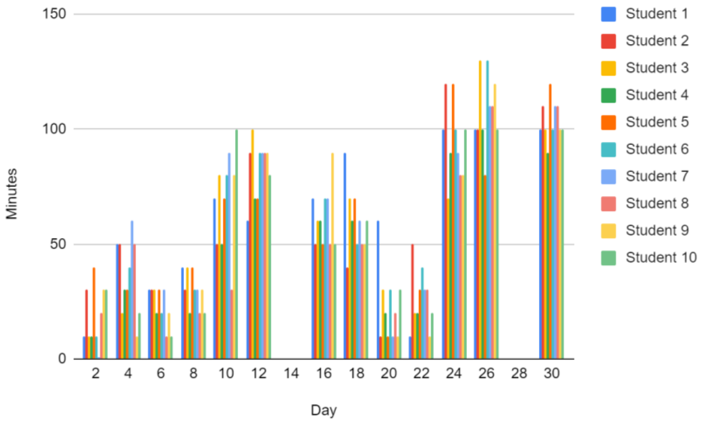


Figure 5. How long a student went outside in minutes every two days

After both experiments were finished, the results seem to solve all of the challenges listed in section 2. Before our product was developed we set out a goal, which was to balance a student's/children's indoor activities with their outdoor activities and limit the amount of nonproductive activity being performed indoors. The results of our experiments show that students are spending more time outside than before, and also went outside more when they didn’t use our product. Even though there are times when a student doesn’t spend much time outside or doesn't go outside much, the average of all ten students shows a positive result.

5. RELATED WORK

Tuan.N. et al summarized the challenges of online learning and the negative, mixed, and null findings of online learning. As a result, the author proved that online learning is not at least as effective as traditional learning [11]. Additionally, the author provided multiple examples of the negative impact on students due to online learning. Related to this paper, we added to their research by proving the negative side of online learning using an accelerometer and a tennis racket. We were able to demonstrate that online learning has many flaws, such as not knowing what a student is doing at all times and the nonproductive activities many students perform. Rather than providing the facts about the flaws of online learning, we were able to solve these problems by creating an application that allows students' parents to see their productivity and also balances students' indoor activities and outdoor activities.

Brian.G. et al presented an application that uses a three-dimensional accelerometer sensor to measure a human hand motion [12]. By using a three-dimensional accelerometer sensor the author was able to track the movement of the hand such as sign language. Instead of using an accelerometer sensor to measure a human hand motion, we coded it into measuring the speed at which a tennis racket is swung to reduce the amount of nonproductive activity being performed by students during online school learning. Our experiment results show that students were able to balance their time spent indoors and their time spent outdoors.

Sian.L. et al explore the topic of movement recognition using the accelerometer in smartphones [13]. By including an accelerometer sensor in a smartphone, they were able to recognize common movements such as walking, standing, sitting, etc. Instead of using an accelerometer sensor in a smartphone to recognize movement, we attached an accelerometer sensor on a tennis racket to recognize the speed at which the racket is swung and also record that data onto a database. Both papers use an accelerometer to track movement to solve a specific problem.

6. CONCLUSIONS

In this project, we proposed software that limits the number of unproductive activities being performed during online learning for students and also balances their indoor activities and outdoor activities by tracking the speed at which a tennis racket moves [14]. A website has also been developed to log all the data collected by the accelerometer that tracked the movement of the tennis racket. The data will then be sent to their guardian through email or phone showing their child's daily performance. The results show that our prediction is correct. After using our product for a couple of days students are able to balance the activity and time they spent indoors with the activity they spent outdoors. Before students tend to spend more time indoors than outdoors but after a few days students with our product they're more active outdoors and seem to be more energized and engaged than before.

As for our product, a limitation related to our application is that it doesn’t show the performance of these tennis players, such as how well they played, what they need to work on such as backhand groundstroke, or a forehand groundstroke, and what steps they should follow to correct it. Another limitation is that our product is only limited to tennis rackets, which could be expanded in the future to different sports that are harder to track, such as basketball or soccer.

In our future updates, we plan on adding new ideas to our website, such as showing a visual demonstrating how the user swung their racket, what they are doing wrong in their tennis games, and what they can do and work on to improve their own inner game of tennis. We also plan on creating a better design for our software/product so that it isn’t limited to just tennis rackets instead it would work on all sports items, such as basketball, soccer ball, etc.

REFERENCES

[1] He, Feng, Yu Deng, and Weina Li. "Coronavirus disease 2019: What we know?." Journal of medical virology 92.7 (2020): 719-725.

[2] Akin, Levent, and Mustafa Gökhan Gözel. "Understanding dynamics of pandemics." Turkish journal of medical sciences 50.9 (2020): 515-519.

[3] Allen, Tom, Simon Choppin, and Duane Knudson. "A review of tennis racket performance parameters." Sports Engineering 19.1 (2016): 1-11.

[4] Peng, Minhua. "Outbreak of COVID-19: An emerging global pandemic threat." Biomedicine & Pharmacotherapy 129 (2020): 110499.

[5] Ganne, Pratyusha, et al. "Digital eye strain epidemic amid COVID-19 pandemic–a cross-sectional survey." Ophthalmic epidemiology 28.4 (2021): 285-292.

[6] Van Rompay, Thomas JL, Dorette J. Vonk, and Marieke L. Fransen. "The eye of the camera: Effects of security cameras on prosocial behavior." Environment and Behavior 41.1 (2009): 60-74.

[7] Baumol, William J. "Entrepreneurship: Productive, unproductive, and destructive." Journal of business venturing 11.1 (1996): 3-22.

[8] Khawas, Chunnu, and Pritam Shah. "Application of firebase in android app development-a study." International Journal of Computer Applications 179.46 (2018): 49-53.

[9] Shaffer, David Williamson, et al. "Video games and the future of learning." Phi delta kappan 87.2 (2005): 105-111.

[10] D'Angelo, John, and Sherry K. Little. "Successful web pages: what are they and do they exist?." Information technology and libraries 17.2 (1998): 71.

[11] Nguyen, Tuan. "The effectiveness of online learning: Beyond no significant difference and future horizons." MERLOT Journal of Online Learning and Teaching 11.2 (2015): 309-319.

[12] Graham, Brian Barkley. Using an accelerometer sensor to measure human hand motion. Diss. Massachusetts Institute of Technology, 2000.

[13] Lau, Sian Lun, and Klaus David. "Movement recognition using the accelerometer in smartphones." 2010 Future Network & Mobile Summit. IEEE, 2010.

[14] Weng, Pei-Yi, and Yen-Cheng Chiang. "Psychological restoration through indoor and outdoor leisure activities." Journal of Leisure Research 46.2 (2014): 203-217.

[15] Pierce, James. "Smart home security cameras and shifting lines of creepiness: A design-led inquiry." Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019.

PREDICTION OF CHRONIC AND NON-CHRONIC KIDNEY DISEASE USING MODIFIED DBN WITH MAP AND REDUCE FRAMEWORK

P. Ravikumaran[1], K. Vimala Devi[2] and K. Valarmathi[3]

[1]Dept. of Computer Science and Engineering, Fatima Michael College of Engg & Tech, Madurai- 625020, Tamil Nadu, India
[2]School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Vellore Campus, Vellore- 632014, India
[3]Dept of ECE, P.S.R Engineering College,
Sivakasi- 626140, Tamil Nadu, India

ABSTRACT

Modern medical information comes in the form of an enormous volume of data that is challenging to maintain using conventional methods. The advancement of big data in the medical and basic healthcare societies is facilitated by precision medical data research, which focuses on comprehending early illness, patient healthcare facilities, and providers. It concentrates primarily on anticipating and discovering direct analysis of some of the substantial health effects that have increased in numerous countries. The existing health industry cannot retrieve detailed information from the chronic disease directory. The advancement of CKD (chronic kidney disease) and the methods used to identify the disease is a difficult task that can lower the cost of diagnosis. In this research, a modified MapReduce and pruning layer-based classification model using the deep belief network (DBN) and the dataset used as CKD were acquired from the UCI repository of machine learning. We have utilized the full potentiality of the DBNs by deploying deep learning methodology to establish better classification of the patient's kidney. Finally, data will be trained and classified using the classification layer and the quality will be compared to the existing method.

.

KEYWORDS

Chronic kidney disease, deep belief neural network, MapReduce, Pruning layer.

1. INTRODUCTION

Now a day's knowledge on data is higher and required to process a very large amount of data and time-consuming. Thus, data mining methods can be utilized to classify patients' diseases and one of the major diseases considered is Kidney disorder [8]. CKD is a disorder when kidneys are destroyed and the disposal fluids produced by the physique stay internally that origins health issues. The primary cause is a disease like diabetics or heart problems with heavy blood pressure [6]. Other risky situations triggering CKD involve heart disease, obesity, and family hereditary chronic disease. It may cause very expensive kidney transplantation. So, treatment should be done earlier. On large-scale aggregated datasets, existing simulation approaches aim to attain inaccurate diagnostic identification [9].

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 185-202, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121615

In recent times, deep learning has received a great deal of prominence from both industry and academia owing to its favourable efficiency in several practical issue [15]. DBN with RBMs are the most essential multi-layer network architectures for deep learning. RBM contains an input and a hidden layer. There have been no neural circuits with same layer and it has a collection of connection weights for input weights as well as hidden neuron. The figure below shows the RBM graphical network, which seems to be a form of energy model.

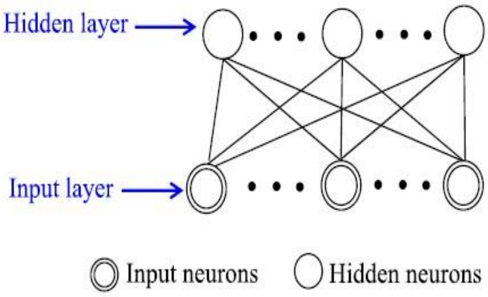


Figure 1. Restricted Boltzmann Machines

DBNs are creative networks trained to retrieve a hierarchical trained feature representation of raw data by optimizing the possibility of the data for training. A single RBM is a key component for deeper architectural design, stacked on top of one another, going to take the previous outcome which is managed as the input following each RBM data type appropriately [11]. Figure 2 shows an explanation of a DBN with stacked RBMs. Label information is used in DBNs to enhance discriminative power for factors learned from the previous RBMs which may not be ideal for variables learned afterwards.

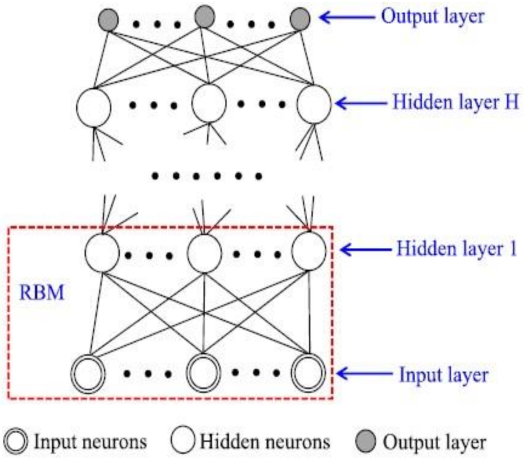


Figure 2. Distributed Belief Network Architecture

The objectives of the current work which are aimed to be achieved are as follows:

- To make effective use of Distributed Belief Network (DBNs) by deploying deep learning.
- To establish better classification of the patient's kidney for protecting them critical conditions.
- To be able to successfully distinguish between non-chronic and chronic kidney diseases with the successful deployment Map Reduce.

To achieve effective results with minimal features along with remarkable connectivity.

2. LITERATURE REVIEW

Jang, Choi et al. [17] developed Deep learning-based radio graphical examination of the hip could predict the osteoporosis. In this development benefit yielded is mere radio graphical examination of hip was sufficient enough to distinguish between the osteoporosis cases and non-osteoporosis without much clinical difficulties

A Temporal modified gradient boost machine was devised by Song, Waitman et al [16]. towards the Longitudinal risk forecasting of the chronic level condition in kidney in patients, who were suffering from diabetes diseases. They utilized data of 14,039 adults, who are suffering type 2 diabetic condition. Higher outcome yielded was ROC of 0.83. However, the pre-defined exclusions in the work could result in wrong false positive values being estimated for diabetic patients with kidney problems, who haven't screened for diabetes readings.

García-Gil et al.,[5] proposed two main pre-processing methods for big data analysis such as homogeneous and heterogeneous ensemble filter to eliminate the noise data. It contains particular prominence in its scalability and performance behavior. This new method is introduced to remove noise data that gives the high quality as well as noise removal clean data also called Smart data. Here, for the experimental evaluation four data sets are used such as HIGGS, Epsilon, SUSY, and ECBDL14. Several stages of class noise have been introduced to examine the cause of deploying such platforms and the development has achieved concerning accuracy using two classifiers like a Knearest neighbor (KNN) and a decision tree (DT) technique. KNN is a noise sensitivity technique if the number of the chosen neighbor is low whereas DT is known as tolerant to noise. Considering the big data issues, the classifier profits from noise treatment even if no additional noise is caused since big data problems involved noise due to incidental homogeneity, accumulation of noisy instances, and spurious correlations. The obtained results showed that noise can be effectively solved in the conceptual methodology. In particular, the homogeneous set is the first effective methodology for attempting to deal with big data noise issues, with low computing moments and encouraging the classifier to achieve greater accuracy.

Abdelaziz et al., [1] Designed an example of cloud-based health services to predict CKD. Cloud computing and also IoT acts an important portion in health care. IoT big data smartphone prediction involves sending large amounts of data to CKD in attempt that would save them in cloud computing. The estimation of harmful diseases like cloud-based CKD –IoT is known to be a big issue facing health care stakeholders. Cloud computing helps patients estimate CKD anyplace at any period in smart cities. The author presented an Intelligent Hybrid Model to Predict CKD-IoT utilizing two intelligent, (linear regression) LR and NN techniques. LR is commonly applied to estimate the major risk factors for CKD. NN can be used to calculate CKD. The results show that when it comes to predicting CKD, the hybrid smart model is 97.8 percent accurate. A hybrid smart model has also been used on Cloud Services as a cloud computing system for predicting CKD in order to assist people in smart cities. The suggested framework is 64% higher than many of the models alluded to in similar works.

Ismail et al., [19] Summarized the implications of big healthcare analytics and relevant applications for both EMR and sensor data. To overcome the issues of irregularity and sparsity in health care data processing algorithms and systems of health care analytics and applications are presented. The proposed outcome depends on an additional layer called middleware which is located between sources of heterogeneous data as well as the Map-reduce Hadoop cluster. The result revealed that the issues that arrived with heterogeneous data have been solved. Furthermore, the availability to incorporate the system with deep learning models that add the capability to identify a particular disease in patients. The proposed framework will be employed in several applications using different optimization methods to minimize the processing time.

Koti & Alamma, [7] The use of health care databases in data analytics, as well as the methods for analysing big data, were discussed. Massive amounts of data are generated in wellbeing organisations with respect to audio, video, text, and (Electronic Health Records) EHR. The big data analytics takes into account the eruption of data to attain significant information that aid throughout the development of effective decisions. After evaluating large data sets of healthcare services in their results, data analysis can improve the operation by using effective methods to achieve the outcomes.

Lokeswari et al., [10] Proposed a new technique designed to improve the above benefits about scalability through maximizing the amount of nodes in the Hadoop cluster and analyzed the efficacy of the classification techniques such as Naïve Bayes, decision tree, and K-nearest neighbor. These parallel methods could be applied in many other biotechnological fields where big dataset forecasting is critical. It offer advantages such as reduction training time, reduced memory requirements, and reduced execution time. There are numerous issues at stake when running parallel algorithms of data mining in cloud systems. It's crucial to split data between processors in a way that minimises computational dependence, communication, proper coordination, overlay interaction, workload averaging among nodes in architectures, and disc IO costs. Running parallel data mining methods upon on Apache Hadoop Map Reduce platform will help with some of these problems. It increases productivity while lowering computational cost, training time, prediction accuracy, and IO availability.

Ahmad et al., [2] A methodology with two major stages, such as categorization modelling and system development, was proposed. Data collection, preparation, grouping, classification, and rule extraction are all part of categorization modelling. Patients with kidney failure have the potential to progress to the chronic stage. A slow decline in kidney function over 3 months characterises CKD, resulting in a severe termination of function (kidney). The goal is to provide such a doctor with a decision-making tool for diagnosing patients with kidney failure. The scheme showed the impact of predicting whether or not due to renal disease have chosen to access the chronic renal progression of the disease. Initially, the processed rules were used to develop the system. This study showed a method that accurately identified a CKD occurrence depends on various factors with a 98.34 percentile accuracy rate. This scheme is utilized to help the physician in accurately measuring the chronic illness of kidney diseases in humans.

Sahoo et al., [13] A probabilistic data collection and correlation analysis of the collected data has been designed. Analysis of the medical fields and expectations about future medical problems are still in the insightful phase. The Data Analysis Framework was also cloud-enabled, and it is the optimal way to analyse structured and unstructured information generated by health-care management solutions. Performance analysis of the proposed processes is carried out and uses extensive simulations that provide 98 percent accuracy in the cloud environment. And retains 90 percent of the CPU as well as bandwidth utilization to shorten the length of analysis. Also, the cloud-based MapReduce model is often used as system architectures for our data analytics. Our method can be used for a variety of health or patient remote monitoring applications, including cardiovascular predictive modelling or tumor severity classification, as per the researchers. The new design will be validated in the realworld healthcare sector using real-time analysis systems such SPARK.

Ed-daoudy & Maalmi, [4] Focused on the application of the distributed machine learning algorithm to stream health data activities absorbed via Kafka concepts to stream processing. First of all, utilizing Spark rather than Hadoop Map Reduce which is restricted to real-time computation, we convert the standard decision tree method into such a parallel, distributed, scalable and fast DT. Secondly, to forecast health status, this method is adopted to transmitting information coming across distributed sources of a different illness. The system predicts health status based on a variety of input data, transmits an warning notification to healthcare professionals, and data is stored in a database system for evaluation of health data and reporting streams. DTs tested the output against data machine learning methods such as WEKA. Finally, to demonstrate the efficiency of the proposed architecture, performance assessment variables like throughput and execution time are measured.

After reviewing all the works reviewed above, the scope and Motivation of this work is, Studies on the precision medical data have focused on understanding early onset of illness as well as the patience health care providers and centre, which had led to the progression of big data in the medical and fundamental healthcare societies. As a result, the concentration on expecting and investigating the direct impact of some substantial health effects have raised in several countries. The existing methodologies in the health sector can’t obtain useful information from the any kind of chronic disease directory. The advancement of CKD (chronic kidney disease) and the methods used to identify the disease had also become a daunting task in lowering the diagnosis cost. Some of the identified research gaps that were identified are as follows:

- Several Deep Neural network-based behavioral systems are still in developmental phase as the understanding of the unique behavior of patient’s chronic conditions need to get improvise a lot.
- One among the necessary functions in medical care is processing of the data, which should be supported by the big data analytics platform. However, many of the traditional algorithms were found ineffective in handling vast data.
- The prediction performance of the model needs to be improved and better concentration on significant feature is required in order to deal with the sensitive data in the medical care.

There were works Comito, Talia et al. [3] ; Comito, Talia et al. [18]; and Comito, Falcone et al. [3] reported in the literature that concerned about the various issues pertaining to the mobile applications like better routing, resource delegation, and management of the energy. However, the current work will only be considering the benchmark dataset, which doesn’t require any data gathering. Thus, deployment of mobile devices was not required and usage of it was not done.

3. PROPOSED METHODOLOGY

The overall flow of the methodology has been explained in the figure 3. Initially, the input dataset has been taken. Here, the used dataset is Chronic Kidney Dataset (CKD). We proposed a novel MR (Map Reduce) and pruning layer-based classification model using the deep belief network (DBN). The novelties of the work carried out are discussed below.

The remarkable connectivity features are selected through a pruning deep belief network algorithm. DBNs are then stacked by plenty of the restricted Boltzmann machines (RBMs). The bottom layer is mainly used to retrieve the input data vector and transfer of the input data to the hidden layer is done through RBM, that is, the output of the lower layer RBM goes to the input of the higher layer RBM. This special case based-energy generation model serves as a learning model for randomly distributed data and then the pruned feature will be given as an input for the DBN’s RBM. In Back propagation, we will incorporate Map and Reducer to handle the huge volume of data.

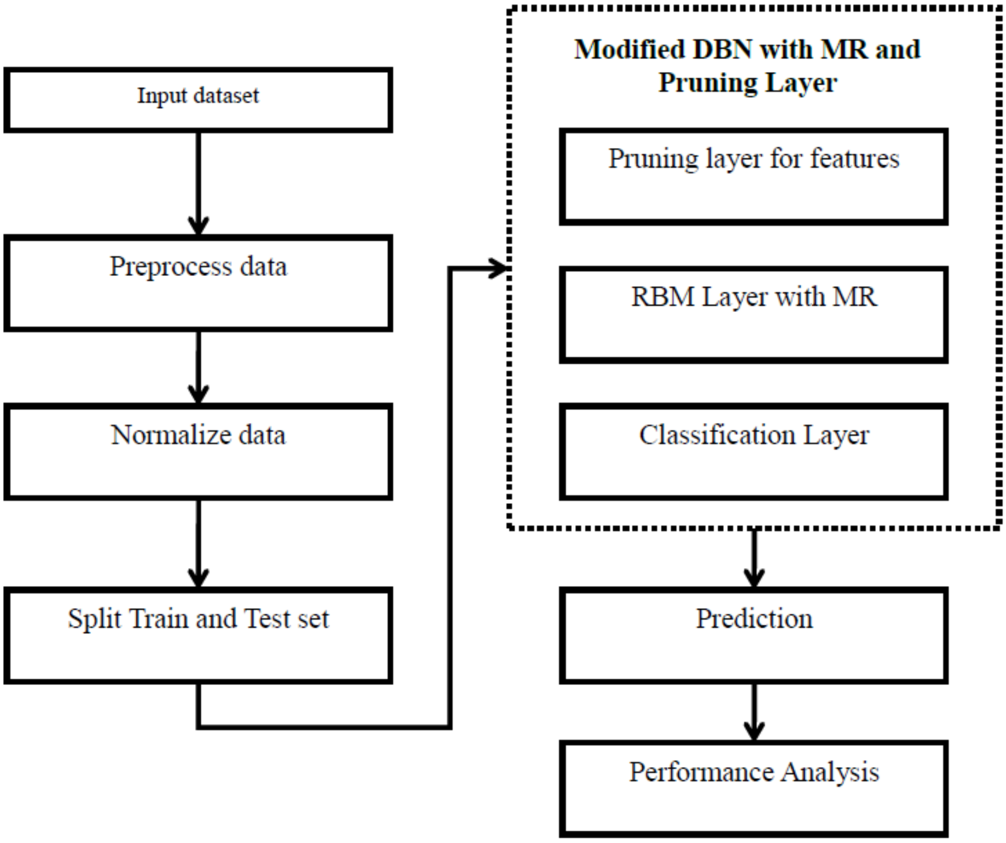


Figure 3. Overall flow of Modified DBN with MR and pruning

The given data are pre-processed and then the data are normalized using standard scalar. It normalizes features of the dataset through scaling to unit variance and this process is very usual in pre-processing step. It also avoids features with large variances from employing a large influence during model training. Then the normalized features are split as training and test set. The next step is feature pruning which took relevant features and sent to the next layer. This creates a DBN layer and that pruned features are sent to the RBM layer with map Reduce. During this process, classification is done in the RBM which predicts the chronic and non-chronic disease, and the performance analysis was done to know the accurate prediction of the given data.

Algorithm 1

Modified Deep belief network (DBN) with MR and pruned layer

Input: Number of epoch e, internal hidden number of layersN_l , hyper-parameters, hidden nodes, v visible node.

Output: The trained DBN model

Step 1: Pruning

```
While $N_l \neq 0$
{
```

```
Initialize the hyper-parameters for the network based on the configuration of the model at the
master node

        Divide the training set into subsets
        Send Parameter information to each worker node from the master node
        Allocate jobs to worker node (Subset transfer) $W_n$
        for i = 1 to $W_n$
{
                for j = 1 to e
{

Calculate Gibbs sampling to determine the estimated w, b, and c gradients
Save the average parameter results to the master node.

}
Transfer the trained network to the master node.
}

Create next layer of RBM

Save the average parameters to global parameters that were calculated from each worker node.

}
```

Step 2: $\theta\ updation$

```
Initialize all hyper parameter values for pruned features and create hidden and visible nodes
for the training dataset.

for i in h:
{
        Execute activation function using equation (7)
}

for j in v:

{
        Execute activation function using equation (8)
}

Execute equation (11, 12, 13) with activation function to return learned w,b, c($\theta$) $value$
```

Step 3: Mapping

```
//Mapping function for each key-value pairs
Execute equation 11, 12, 13 to calculate the initial $\theta$ value
```

Step 4: Reducing

```
//Reducing function for each key-value pairs
Execute step 2 to update $\theta$ value.
```

Compute mean squared error (MSE)
Store all MSE and corresponding θ value

Step 5: Prediction
Test data with the trained model
Return predicted data.

3.1. Pruning

Pruning is the first step in emerging distributed architecture. That master node sets up a network specification and parameters with the total number of neurons inside the visible and hidden layers, prejudice of the input and hidden layers, amount of epoch, and the learning rate, etc. The dataset is then divided into several subsets at random and broadcasted to all staff nodes, including a copy of all specified parameters. Based on the training setup, the amount of spits is measured dynamically. Afterward, each task requires the sampling methodology (Gibbs) to evaluate the estimated gradient w, b, and c over his or her portion of the break for each epoch. We regularly measure the gathered parameters for each task and modify the global parameter configuration, which is modified w, b, and c, as well as their cumulative estimated gradient, in the intermediary storage approach. The masters acquire a copy of the training set and initialize this for another layer of RBM because after training is completed. The training of the resting stage of RBM is close to that of the lowermost stage RBM except the input is modified. The modified DBM is developed using the spark algorithm from algorithm 1. The distribution of the learning process is achieved using the data-parallel method, which is focused on the methodology suggested above. We keep a copy of the original model on each workforce machine, and then method various training subsets for each worker. Depending on the synchronous variable averaging method, the effects are averaged and the model parameters are synchronized.

3.2. Map and Reduce Function

MapReduce provides a scheduling algorithm for performing distributed computing on multiple computers.

The system administrator modifies a Master Controller procedure, a sequence of mappers and reducers activities on multiple systems in the scheme. One MapReduce job is computation, which comprises of two stages involving the map and reduces operations. The map principle specifies how the input is segmented into a sequence of subset that are divided into pairs and assigned to mappers. In this regard, each mapper employs the user-specific mapping designed for each input pair and outputs a set of middle pairs that are indicated to the map computer systems' local discs. The source code sends these middle pairs to the master, who is in charge of informing the reducers about these locations.

When the reducers read all of the transitional pairs in digital format, they organise and assemble the keys. To manage all values from each individual key and create a modern significance with seperate key, each reducer uses a user-defined reduce mechanism. All reducers' resulting key-value pairs are acquired as findings and transferred to the output file. Finally, the system of MapReduce completes all of the challenges in comparison. The MapReduce framework can thus be used for highlevel segmentation as well as data processing.

3.3. Probabilistic graph framework:

It is a type of generative stochastic computer program which an understand a probability distribution throughout its inputs. It is made up of both visible and hidden nodes. A RBM is additionally constrained by the elimination of visible-visible and hidden-hidden links. Figure 1 depicts a graphical representation of an RBM. Its probability is described as,

$$p(k,h,\theta) = \frac{e^{-E(k,h,\theta)}}{Q} \qquad (1)$$

$$Q = \sum_{\hat{k}} \sum_{\hat{h}} e^{-E(\hat{k},\hat{h},\theta)} \qquad (2)$$

where $e^{-E(k,h,\theta)}$ signifies the energy function
θdenotes the parameters
Qindicates the normalizing factor and also known as the partition function
$k, h$is the two vector variables denoting visible and hidden nodes

The energy function is computed as follows

$$E(k,h,\theta) = -b^T k - c^T h - h^T wk \qquad (3)$$

Where $b_{j\, and}$ c_j are the offsets and w_{ij} denotes the connection weight.
The likelihood $p(k)$denotes the probability allotted to visible vector x, which is added overall feasible configuration of hidden nodes that is

$$p(k) = \sum_{\hat{h}} p(k,\hat{h}) = \sum_{\hat{h}} \frac{e^{-E(k,\hat{h})}}{Q} \qquad (4)$$

Negative log-likelihood is the optimal solution, which is calculated as,

$$L_l(\theta, Tr_s) = -\sum_{k\epsilon Tr_s} \log P(k,\theta) \qquad (5)$$

Where $\theta = \{b,c,w\}$ and Tr_s is the training dataset. According to Bayesian statistics, the issue is to estimate the parameters governing the model through minimizing the negative log-likelihood $\log P(k)$ that is

$$\min_{\theta} L_l(\theta, Tr_s) \qquad (6)$$

The comprehensive procedure for solving the dual issue of maximum likelihood that used a stochastic gradient descent framework will be reviewed in the following paragraph.

3.4. RBM Learning Method

Based on the specific structure of the RBM, visible as well as hidden nodes are linearly separable from each other. Where k_j and $h_i \in \{0,1\}$are the probabilistic versions of the regular neuron kernel function, the probabilistic version of the regular neuron activation functions is provided by

$$p(h_i = 1 \,|\, k) = \frac{e^{c_i + w_i k}}{1 + e^{c_i + w_i k}} = L_{sig}(c_i + w_i k) \qquad (7)$$

$$p(h_i = 1 \mid k) = \frac{e^{c_i + w_i k}}{1 + e^{c_i + w_i k}} = L_{sig}(c_i + w_i k) \tag{7}$$

$$p(k_i = 1 \mid h) = \frac{e^{b_i + w_i^T k}}{1 + e^{b_i + w_i^T k}} = L_{sig}(e^{b_i + w_i^T k}) \tag{8}$$

Where w_j is j-th column of w, L_{sig} is the logistics sigmoid function defined by

$$L_{sig}(k) = \frac{e^k}{1+e^k} = \frac{1}{1+e^{-k}} \tag{9}$$

$$-\frac{L_{sig}\, logp(k)}{L_{sig}\; w} = \exp_{ed}[k.h] - \exp_{\ddot{ed}}[k.h],$$
$$-\frac{L_{sig}\, logp(k)}{L_{sig}\; w} = \exp_{ed}[k] - \exp_{\ddot{ed}}[k],$$
$$-\frac{L_{sig}\, logp(k)}{L_{sig}\; w} = \exp_{ed}[h] - \exp_{\ddot{ed}}[h], \tag{10}$$

Where $\exp_{\ddot{ed}}$ is the expectation over k below the experiential distribution $\ddot{ed}$and is $\exp_{ed}$ the expectation below the distribution of model. It is usually tough to compute this gradient function, though $\exp_{\ddot{ed}}[k.h]$ can be computed easily. The evaluation of $\exp_{ed}[k.h]$ is more difficult, as the actual expectation over all feasible configurations of the input x is costly to evaluate. In most cases, the expectations are calculated using a set amount of unbiased specimens. As shown in the equation, unbiased illustrations of $\exp_{ed}[k.h]$ can be obtained by computing modified Gibbs sampling, where k^0 is a training example from the training dataset, also referred as the order level in Gibbs sampling. The visible and hidden node vectors generated after n-th Gibbs sampling are k^n and h^n, accordingly.

Gibbs sampling is time-consuming. CD (contrastive divergence) learning was adopted as an effective learning method. Two elements of CD learning were used to accelerate the sampling process. After single n-steps of Gibbs sampling, initialise the Markov chain with both the training and obtained sample. The modified conditions for every parameter are given through,

$$w = \epsilon(k^0.h^0 - k^1.h^1) \tag{11}$$
$$\text{b}=\epsilon(k^0.k^1) \tag{12}$$
$$\text{c}=\epsilon(h^0.h^1) \tag{13}$$

where ϵ describes the learning rate. The training issued of pseudo-code for RBM is demonstrated in algorithm 1 after describing the above discussion. Finally, data will be trained and classified using the classification layer and the quality will be compared to the existing method.

4. EXPERIMENTAL RESULTS

4.1. Dataset Description

This study made use of a dataset called CKD, which was uploaded to the UCI machine learning repository in 2015. This dataset contains 25 attributes, 14 of which are nominal and 11 of which

are numerical. [8] The characteristics and their description of the features that cope with our research which is outlined. A maximum inst
ance of the dataset will be used for prediction model development of which has been labeled for chronic and non-chronic kidney disease.

4.2. Performance Metrics

The performance metrics are described below for the proposed system to measure the efficiency of the given research [12].

TP (**True positive**): The province where several instances are categorized as precise as true.

FP (**False positive**): The entity in which the amount of scenarios is assembled as precise as they were not correct.

FP (**False Negative**): The province in which the number correctly classified is categorized as false as being probably true.

TN (**True negative**): The situation that the amount of data is classified as false because they were untrue.

Accuracy:

It is a metric of the computation sector of the formulation that represents the systematised error. The difference among the potential outcome and the real value is often caused by low accuracy. This means that the machine evaluates the exceptional input variables several times using the same procedure, and the results are consistent. The amount of real outcomes in the total is known as accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (14)$$

Precision:

It's a metric of algebraic variance that describes random error.

$$Precision = \frac{TP}{TP+FP} \quad (15)$$

Sensitivity:

It is known as the true optimistic rate, recall, or detection possibility in some fields, is a metric that measures the percentage of true positives that are correctly identified [9]

$$Sensitivity = \frac{TP}{TP+FN} \quad (16)$$

Specificity:

In some fields, sensitivity is named as the true positive rate, recall, or probability of detection, measures the percentage of actual positive that is expected.

$$Specificity = \frac{TP}{TP+FP} \quad (17)$$

Recall:

It is the total of all true positives in the positive category.

$$Recall = TP + FN \tag{18}$$

Root Mean Square Error RMSE: It is one of the most widely used types for assessing the performance of predictions. It demonstrates how far assumptions fall from the true value measured using the Euclidean distance. RMSE is also used in supervised learning applications as RMSE utilizes and requires true measurement at each predicted data point. RMSE is computed as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y\ (i) - y\ (I))^2}{N}} \tag{19}$$

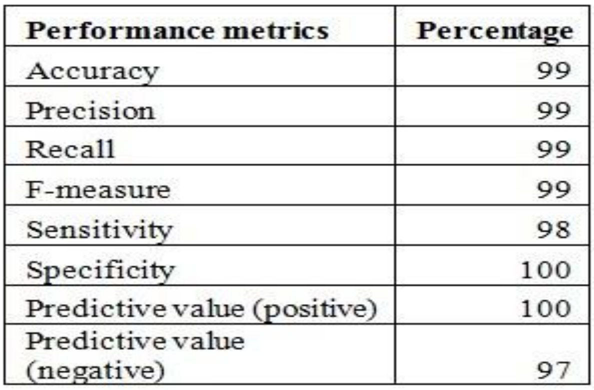
where, $y\ (l)$ indicates the true value and $y\ (i)$ represents the prediction value.

$$\text{Predictive value (positive)} = \frac{TP}{TP+FN} \tag{20}$$

$$\text{Predictive value (negative)} = \frac{TN}{FN+TN} \tag{21}$$

The performance metrics are described with the formula which is used for the simulation performance. The comparative analysis results as explained as below
.

Table 1. Comparison of proposed method in terms of performance metrics



Performance metrics	Percentage
Accuracy	99
Precision	99
Recall	99
F-measure	99
Sensitivity	98
Specificity	100
Predictive value (positive)	100
Predictive value (negative)	97

Table 1 depicts the values of accuracy, precision, recall, f-measure, sensitivity, and specificity of the proposed work which shows the maximum value for the prediction of chronic kidney disease.

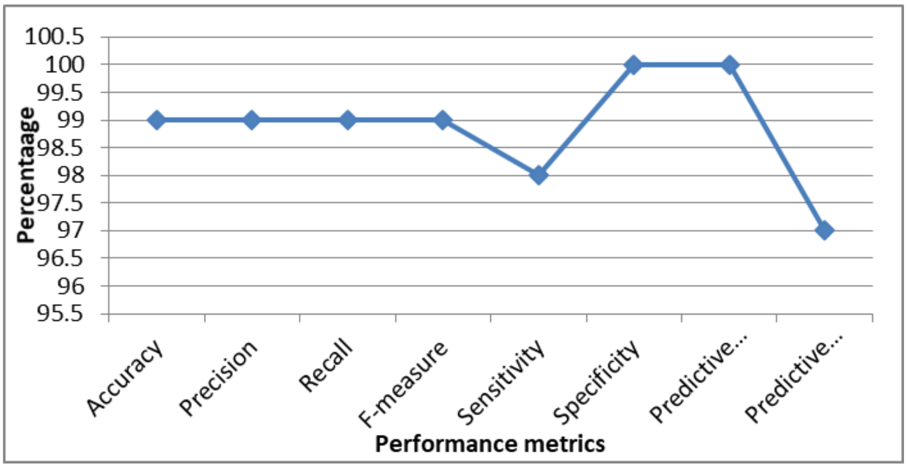
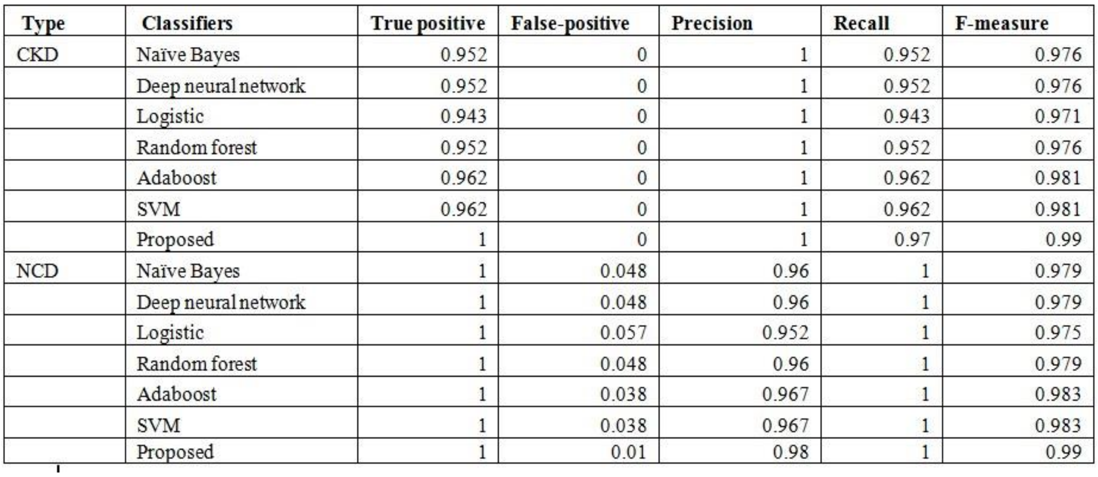


Figure 4. Comparative analysis of the proposed system

Figure 4 depicts the value of all of the proposed performance metrics. The x-axis depicts performance metrics, while the y axis depicts the percentage.

Table 2. Comparison of existing classifiers and proposed system with the type of disease prediction [8]



Type	Classifiers	True positive	False-positive	Precision	Recall	F-measure
CKD	Naïve Bayes	0.952	0	1	0.952	0.976
	Deep neural network	0.952	0	1	0.952	0.976
	Logistic	0.943	0	1	0.943	0.971
	Random forest	0.952	0	1	0.952	0.976
	Adaboost	0.962	0	1	0.962	0.981
	SVM	0.962	0	1	0.962	0.981
	Proposed	1	0	1	0.97	0.99
NCD	Naïve Bayes	1	0.048	0.96	1	0.979
	Deep neural network	1	0.048	0.96	1	0.979
	Logistic	1	0.057	0.952	1	0.975
	Random forest	1	0.048	0.96	1	0.979
	Adaboost	1	0.038	0.967	1	0.983
	SVM	1	0.038	0.967	1	0.983
	Proposed	1	0.01	0.98	1	0.99

Table.2 the quantity of the true positive rate is calculated as the median of all positive values within each epoch. For each technique, the positive result rate is computed in the same way. Precision indicates how many tuples' metrics the classifier classified as chronic kidney disease using structure we developed. For F-measure, the harmonic mean of accuracy and recall is well founded. The designed methodology has the greatest accuracy, recall, and true positive valuation of any systems available in the market.

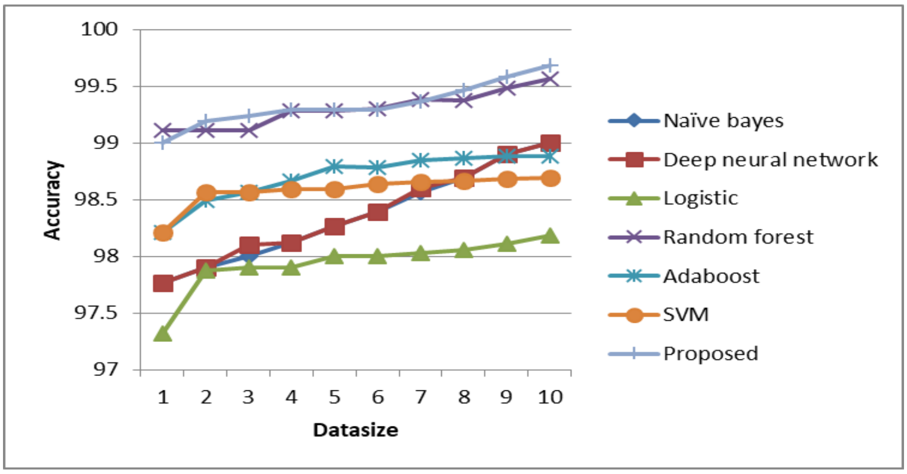


Figure 5. Comparative analysis of accuracy for proposed and existing models

Figure 5 depicts the accuracy measures for the proposed as well as previous methodologies. The x-axis signifies size of data (gigabytes), and the y-axis reflects accuracy (percentage). The results revealed that the proposed method outperforms other methods with the better result of accuracy.

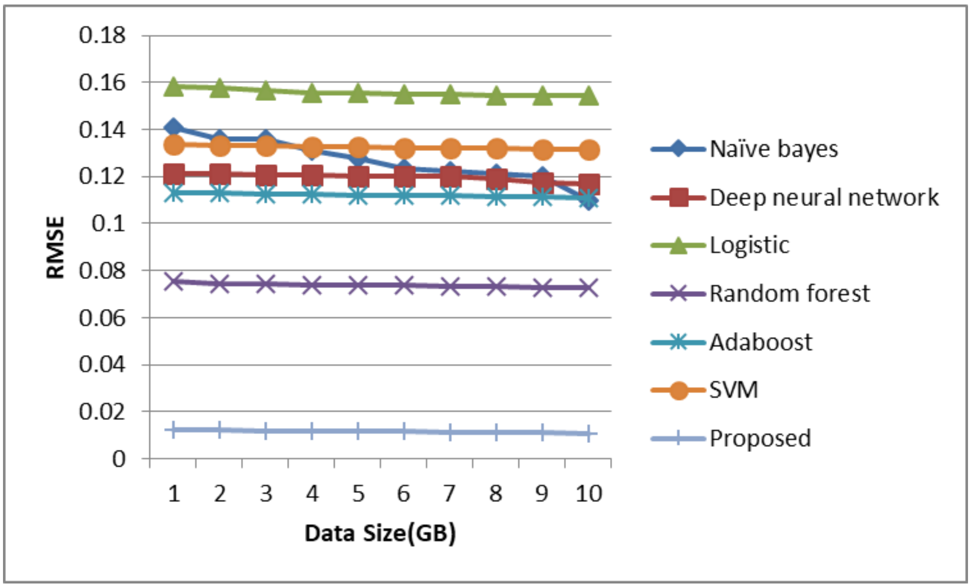


Figure 6. Comparisons of RMSE value

Figure 6 presents a comparison of proposed and current RMSE value techniques. The data size is represented on the x-axis, and the RMSE value is depicted on the y-axis. The results demonstrate that the systems algorithms are superior; the RMSE value is low, indicating better performance.

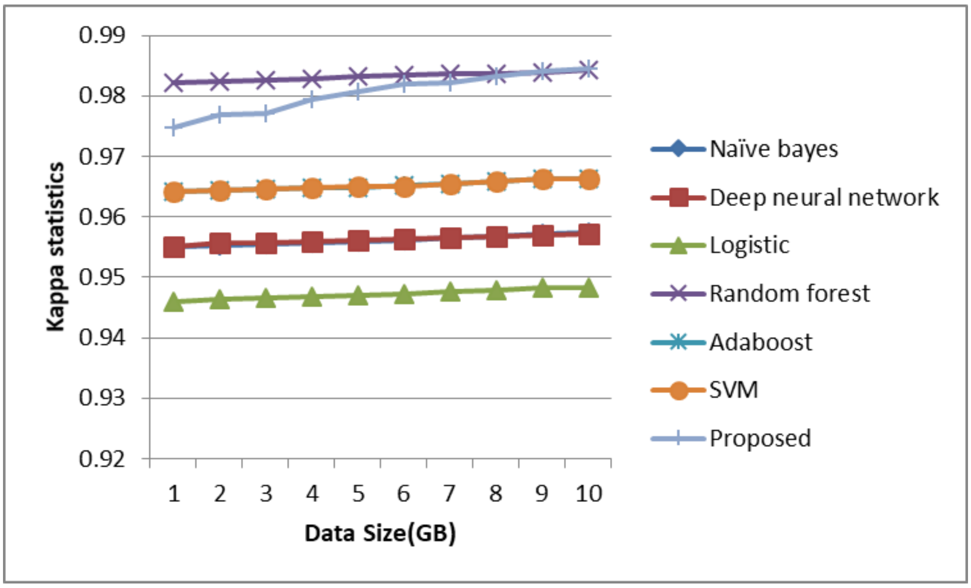
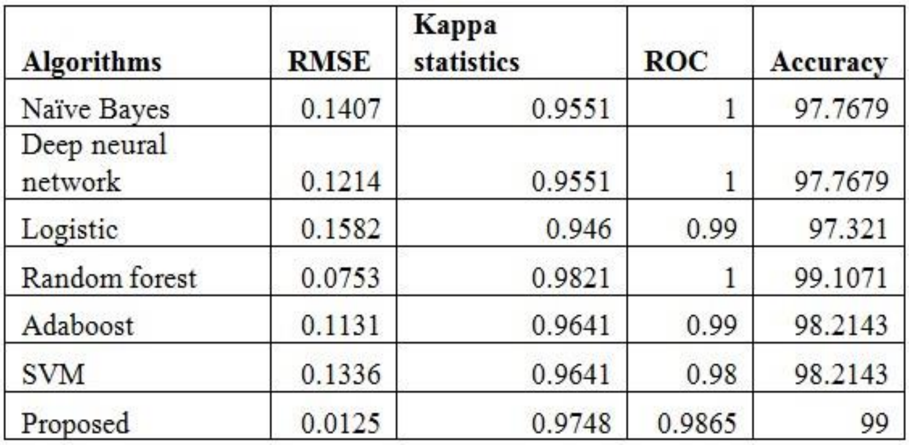


Figure 7. Comparative analysis of Kappa Statistics

Figure 7 represents a kappa statistics-based comparative survey of various methods. The data size is defined by the x-axis, and the value of kappa is indicated by the y-axis. In this case, the proposed kappa value rises as data size increases, whereas the random forest decreases as data size rises.

Table 3. Comparison of proposed and existing methods of RMSE, ROC, and kappa statistics



Algorithms	RMSE	Kappa statistics	ROC	Accuracy
Naïve Bayes	0.1407	0.9551	1	97.7679
Deep neural network	0.1214	0.9551	1	97.7679
Logistic	0.1582	0.946	0.99	97.321
Random forest	0.0753	0.9821	1	99.1071
Adaboost	0.1131	0.9641	0.99	98.2143
SVM	0.1336	0.9641	0.98	98.2143
Proposed	0.0125	0.9748	0.9865	99

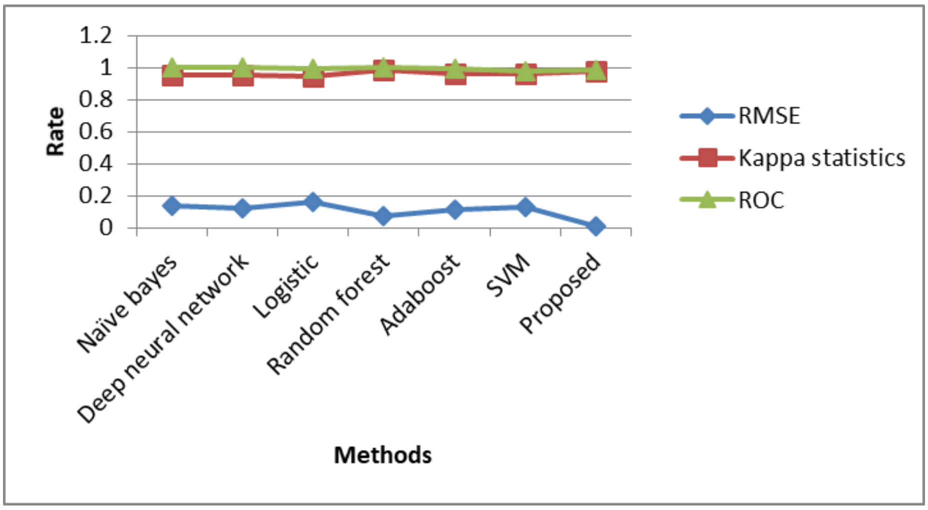


Figure 8. Comparison of RMSE, ROC, and kappa statistics

Table.3 and Figure 8 show that as the amount of epochs increases, method loss decreases. The training set's loss function is set to zero, and the test set's loss function is slightly greater than zero. The RMSE deals with the data's outcome, supplying a severity despite its negative and positive standard errors described in the prediction methods. Kappa statistics show the true percentage of raters who agree on the prediction of CKD with a score of 0.97, indicating nearly perfect agreement. On the ROC curve, the area under plot curve represents false positive versus true positive. We might have had the entire distance taken up by the curve, so the valuation for our proposed system is 98 percent and the precision is 99 percent, indicating a higher level of accuracy.

5. Conclusion

In this paper, a modified DBN with MR and pruning layer has been proposed to manage the massive dataset samples. The methodology utilizes a Map-Reduce algorithm incorporated with a modified DBN to receive the predicted outcomes efficiently with trained persistent DBN with CKD datasets. The standard scalar approach is used for normalizing the data in the preprocessing steps to predict CKD. This proposed methodology predicts chronic and non-chronic kidney disease. Thus, the combination of the modified DBN classification with the MR platform is proficient for giving better outcomes. The efficacy was calculated in terms of accuracy, recall, sensitivity, specificity, ROC, kappa statistics, and also RMSE which are very crucial in the medical field. Our proposed methodology gave rise to the following advantages, which was able to effectively utilize the DBN by achieving efficient classification performance characteristics as compared in the above sections with minimal and remarkable connectivity between the features to distinguish between the non-chronic as well as the chronic kidney diseases in human beings. Our method was also found to hold good in terms of considered big data. The limitations that our work had was that deploying the benchmark dataset instead of data gathering, which could have yielded more practical applicability. Furthermore, we also desire to make use of various mobile devices by also taking into the consideration various challenges pertaining to the mobile data collection operations.

REFERENCES

[1] Abdelaziz, A., Salama, A. S., Riad, A. M., & Mahmoud, A. N. (2019): A Machine Learning Model for Predicting of Chronic Kidney Disease Based Internet of Things and Cloud Computing in Smart Cities. https://doi.org/10.1007/978-3-030-01560-2_5 pp. 93–114

[2] AHMAD, M., TUNDJUNGSARI, V., WIDIANTI, D., AMALIA, P., & RACHMAWATI, U. A. (2022): FUZZY LOGIC-BASED SYSTEMS FOR THE DIAGNOSIS OF CHRONIC KIDNEY DISEASE. DOI: 10.1155/2022/2653665[3]

[3] Comito, C., D. Talia and P. Trunfio (2011). An energy-aware clustering scheme for mobile applications. 2011 IEEE 11th International Conference on Computer and Information Technology, IEEE.

[4] Ed-daoudy, A., & Maalmi, K. (2019): A new Internet of Things architecture for real-time prediction of various diseases using machine learning on big data environment. Journal of Big Data, 6(1). https://doi.org/10.1186/s40537-019-0271-7

[5] García-Gil, D., Luengo, J., García, S., & Herrera, F. (2019): Enabling Smart Data: Noise filtering in Big Data classification. Information Sciences, 479, 135–152. https://doi.org/10.1016/j.ins.2018.12.002

[6] Khamparia, A., Saini, G., Pandey, B., Tiwari, S., Gupta, D., & Khanna, A. (2020). KDSAE: Chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network. Multimedia Tools and Applications, , 35425–35440. https://doi.org/10.1007/s11042-019-07839-z pp (47–48)

[7] Koti, M. S., & Alamma, B. H. (2019): Predictive analytics techniques using big data for healthcare databases. Smart Innovation, Systems and Technologies, 105, https://doi.org/10.1007/978-98113-1927-3_71, pp 679–686.

[8] Kriplani, H., Patel, B., & Roy, S. (2019): Prediction of chronic kidney diseases using deep artificial neural network technique. In Lecture Notes in Computational Vision and Biomechanics. Springer Netherlands. https://doi.org/10.1007/978-3-030-04061-1_18. pp. 179–187

[9] Larson, E. (1991): Medicare: A Strategy for Quality Assurance. In Journal of Nursing Care Quality (Vol. 5, Issue 4). https://doi.org/10.1097/00001786-199107000-00013

[10] Lokeswari, Y. V., Jacob, S. G., & Ramadoss, R. (2019): Parallel Prediction Algorithms for Heterogeneous Data: A Case Study with Real-Time Big Datasets. Advances in Intelligent Systems and Computing,. https://doi.org/10.1007/978-981-13-1882-5_46. pp 529–538

[11] Merzenich, M. M., Nahum, M., & Van Vleet, T. M. (2013): Neuroplasticity: introduction. Progress in Brain Research, https://doi.org/10.1016/B978-0-444-63327-9.10000-1,pp 14–36

[12] Ramani, R., Vimala Devi, K., & Ruba Soundar, K. (2020): MapReduce-based big data framework using modified artificial neural network classifier for diabetic chronic disease prediction. Soft Computing,. https://doi.org/10.1007/s00500-020-04943-3, pp 16335–16345

[13] Sahoo, P. K., Mohapatra, S. K., & Wu, S. L. (2016): Analyzing Healthcare Big Data with Prediction for Future Health Condition. IEEE Access,. https://doi.org/10.1109/ACCESS.2016.2647619, pp 9786–9799

[14] Wang, Y., Pan, Z., Yuan, X., Yang, C., & Gui, W. (2020): A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network. ISA Transactions, https://doi.org/10.1016/j.isatra.2019.07.001, pp 457–467.

[15] Zhang, K., & Chen, X. W. (2014): Large-scale deep belief nets with mapreduce. IEEE Access, 2,. https://doi.org/10.1109/ACCESS.2014.2319813. pp 395–403

[16] Song, X., L. R. Waitman, S. Alan, D. C. Robbins, Y. Hu and M. J. J. m. i. Liu (2020). "Longitudinal risk prediction of chronic kidney disease in diabetic patients using a temporal-enhanced gradient boosting machine: retrospective cohort study." 8(1): e15510.

[17] Jang, R., J. H. Choi, N. Kim, J. S. Chang, P. W. Yoon and C.-H. J. S. r. Kim (2021). "Prediction of osteoporosis from simple hip radiography using deep learning algorithm." 11(1): 1-9.

[18] Comito, C., D. J. P. Talia and M. Computing (2017). "Energy consumption of data mining algorithms on mobile phones: Evaluation and prediction." 42: 248-264

[19] Ismail, A., Shehab, A., & El-Henawy, I. M. (2019): Healthcare Analysis in Smart Big Data Analytics: Reviews, Challenges and Recommendations. https://doi.org/10.1007/978-3-030-01560-2_2, pp. 27–45

AUTHORS

P. Ravikumaran, An active teacher and researcher scholar of Dr. K. Vimala Devi affiliated to Anna University, Chennai, Tamil Nadu. Having 17 years of experience in teaching out of which 7 years in research. Published about 15 research papers in International/National Journals and conferences

Dr. K. Vimala Devi, An active teacher and researcher. Having above 25 years of experience in teaching out of which 15 years in research. Guiding 10 research scholars for Ph. D in the areas of Computer Networks, Network Security, Text mining, Image Processing , Software Testing, Cloud Computing and Big Data. Published about 120 research papers in International/National Journals and conferences, out of which 17 journal papers indexed in ACM portal and 30 indexed in Scopus and 15 IEEE conf. publications. Reviewer in IEEE Communications letter, IJCS and IJNM of John Wiley publications.

Dr. K. Valarmathi, An active teacher and researcher. Having above 25 years of experience in teaching out of which 15 years in research. Guiding 10 research scholars for Ph. D in the areas of Computer Networks, Genetic algorithims, Neural Networks, Fuzzy logic Security, Text mining, Image Processing , Cloud Computing and Big Data. Published about 114 research papers in International/National Journals and conferences.

SafeLanding: An Intelligent Airbag System for Automated Fall Detection and Protection using Machine Learning and Internet-of-Things (IoT)

Richard Lin[1] and Yu Sun[2]

[1]Phillips Academy Andover, 180 Main St, Andover, MA 01810
[2]California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

Abstract

In recent years, the world has been undergoing a drastic change in its age demographic due to an off balance caused by decreasing birth rates and an increase in the elderly population [1]. While 8.5% of the global population were elders in 2015, studies show that this number will hit 17% by 2050. This project will focus on the efficiency of automatic fall detection and contribute to the evolution of fall protection [2], both within elders and the general population. Through our conducted work, we have developed a wearable device capable of efficient fall detection and transmission.

Keywords

Machine Learning, Fall detection, Mobile APP.

1. Introduction

For such a growing demographic, one of the leading causes of injury and death is falling. The CDC estimates that 36 million falls happen each year among elders, and falling accounts for 95% of hip fractures. Fall mortality rates grew around 30% from 2008 to 2018 and steadily remain on the rise. Therefore, reliable and efficient fall detection is crucial to be developed [3]. While falling presents itself as a persistent issue, adequate steps have not been taken to prevent these threats and integrate efficient solutions into everyday life. Fall prevention devices and aids are not frequently used by elder populations, yet they make a substantial difference in the current amount of harm done and lives lost. There has been a growing incentive to keep elders active, so developing these devices is necessary.

Many existing techniques for fall detection utilize accelerometers to measure the acceleration of the user. However, these proposals falsely assume that acceleration is the only dimension of a person's falling motion [6]. Therefore, these systems may mistake other normal activities (such as running, jumping, or sitting) as falls. Most methods do not pull from an adequate number of sources of information, at times lacking the potentially useful implementation of gyroscopic sensors (measuring angular acceleration), magnetometers (measuring magnetic field strength and orientation), and GPS data to name a few.

David C. Wyld et al. (Eds): AIFZ, SIGPRO, NWCOM, CLSB, MLNLP, DTMN, BDIOT, CSITY - 2022
pp. 203-217, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121616

Furthermore, some existing methods use inefficient means for detection. A study in 2014, although only a proof of concept, uses a smartphone sensor placed in a shirt pocket. This technique may be impractical for a user to replicate, and the data collected may be inaccurate due to the unstable placement of the phone.

Other studies occasionally rely on tracking the final body position of the user in order to register a fall in the system. A project back in 2009 first defines falls as the “unintentional transition to the lying posture,” where the intention is determined by accelerometer and gyroscopic data. While this can be used as a secondary step to verify a fall, systems that rely on this as a first step detection are slower and more complex to process. By requiring the collection of the user’s start and end positions in order to complete the first stage of fall detection, these processes are inherently performing after the fall has already taken place.

Finally, there are not enough efforts that are made towards the user experience of fall detection/prevention. Current fall prevention devices are primarily bulky, unaesthetic, expensive, and/or difficult to operate and reuse. We hope to start bridging this disconnect in this paper.

In this paper, we utilize and manipulate accelerometer and gyroscopic data to detect falls, as this combination of data provides a balance between simplicity and accuracy. Our system follows the similar lines of research, in that the fall is detected when the sensor surpasses a certain value threshold that is abnormal for other human behaviors.

Our process has a number of beneficial features. First, an efficient sensor is used, as the coin-sized device is placed discreetly on the front right hip. The placement of the sensor is used to streamline the data collection process and prioritizes the user’s comfortability. A number of existing research requires complex sensor systems that provide minimal to no gain in accuracy and/or functionality. Furthermore, we created an app (paired to the sensor) that is directed towards elders’ primary caretakers, which includes features such as a live detection display and a fall history log. Caretakers are notified when the user falls, and are able to dial emergency services only if needed. Therefore, we believe our method prioritizes efficiency and minimalism while remaining accurate.

In two scenarios, we demonstrate how the above combination of techniques benefits the efficiency of the device. First, we show the merit to our approach by conducting an extensive number of falls after developing our algorithm. Second, we interviewed a collection of seniors that provided feedback on what features would be useful to have on the app.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we faced during our algorithm experimentation and app integration; Section 3 outlines our solutions to the corresponding challenges mentioned prior; Section 4 presents the relevant experiments and the field interviews we conducted, followed by an acknowledgement of related and noteworthy works in Section 5. Finally, Section 6 gives concluding remarks and discusses future areas of expansion for our project.

2. CHALLENGES

Here’s a brief overview of the common concerns we have identified that one may face in a fall situation.

2.1. User capabilities

Although falling is a potentially life-threatening action, user functionality is still crucial to any system looking to be developed into a product. When elders (or anyone) fall, they are often not able to recover on their own from their position on the ground. Therefore, most elders (who are in need of fall detection the most) have a caretaker or primary contact through which they are monitored for their safety [5]. In this system, caretakers are then the ones responsible for caring for their subject(s), whether in the case of a fall or other injuries. It is crucial for these people to be notified if a fall were to occur so that they may take the necessary further steps to ensure the safety of their subjects. In this system, since fall detection data needs to be processed and updated quickly, simplified models with fewer amounts of steps should be used.

Furthermore, tracking the history of falls of a user can be beneficial for medical practitioners, as it may reveal patterns to a user's behavior that could lead to properly diagnosing medical issues.

2.2. Device functionality and experience

For the device, fall detection devices should prioritize the wearer's experience. Some devices on the market are bulky on the body and heavy to wear [7]. Others are invasive to the user's daily activities, placed on parts of the body that may inhibit motion or cause discomfort. An optimal device is one that is lightweight and relatively small, as there would be little to no drawbacks in attaching such a device to the user.

On top of this, the device should efficiently be connected to the main controller app. The caretaker app should only display the necessary information needed to deduce the best course of action, such as the time of the most recent fall.

2.3. Affordability and accuracy of equipment

Although fall detection may be a tool that saves many lives, that does not mean it has to require expensive technology. While some state-of-the-art devices detect falls reliably, such as the apple watch, they boast a high price point due to its combination of other features. Affordable sensors are more than capable (in build and technology) of gathering accurate fall data. Some previous research has also required the usage of multiple sensors positioned on various points on the body. However, the results provide no substantial increase in accuracy as compared to single-sensor systems.

3. SOLUTION

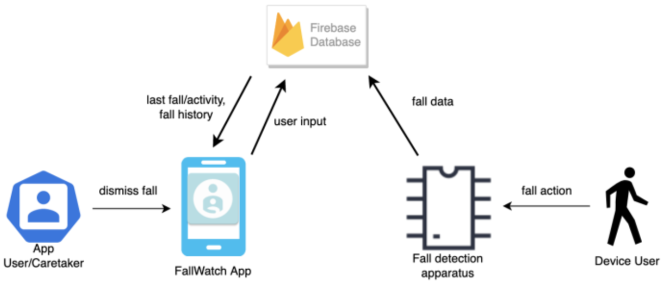


Figure 1. Overview of the solution

FallWatch is an environment that consists of two primary functionalities: the fall detection apparatus and the paired app. The apparatus is a junction of a QT PY ESP32-S3 controller chip linked to a 9-DoF (Degree of Freedom) board, allowing high-quality motion detection. On the board are the LSM6DSOX, the accelerometer and gyroscope sensor, and the LIS3MDL, a magnetometer. The accelerometer and gyroscope sensors have a poll rate of up to 6.7KHz, more than capable of updating the instantaneous data needed in fall scenarios. The accelerometer can measure up to +-16g of force, while the gyroscope +-2000dps (degrees per second).

The environment is built for efficiency. Once a fall is noted from the apparatus through the detection algorithm, the board updates a real-time database [8]. On the other hand, the FallWatch app is also connected to the database and is updated upon a fall. App users then have the option to dismiss and acknowledge the fall, however the sensors still remain active if the user chooses not to do so.

3.2.1 Hardware and Sensors

- Materials and component
- Connection diagrams
- Coding IDE and environment

3.2.2 Detection Algorithm

- Overview
- Diagram
- Code excerpt
- Explanation

3.2.3 Database and Communication

- Firebase introduction
- Pushing the data to Firebase
- Code excerpt

3.2.4 Mobile App

- For each Screen:
- Screenshot
- Code excerpt
- Explanation

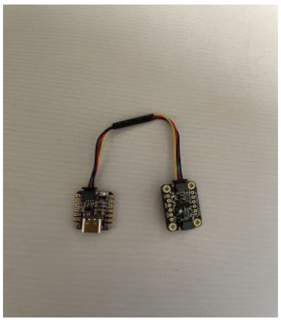


Figure 2. Image of device (sensor and controller boards)

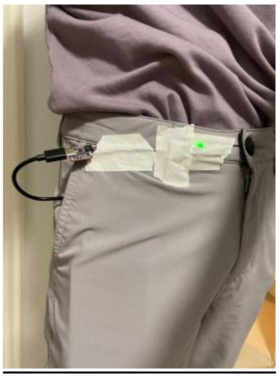
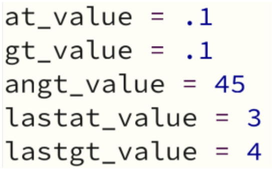


Figure 3. Image of device test position

The controller board (left) and the sensor board (right) are connected via an i2c cable. The dual device is then powered through the controller's USB-C port (with a battery pack or some other source). Both the device and the app are connected to the database over Wi-Fi connection. The sensor is then attached right side up to the hip on the right side of the front of the user's body, in the location where a belt would be worn.

Detection Algorithm [9]

Our system tracks the linear and angular acceleration to process the event of a fall. There are five values that go are used in the algorithm, each representing a threshold for a fall to be detected.

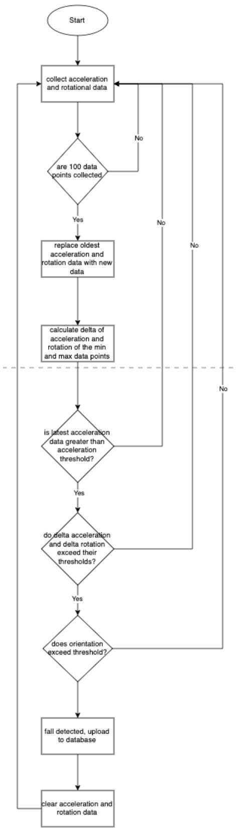


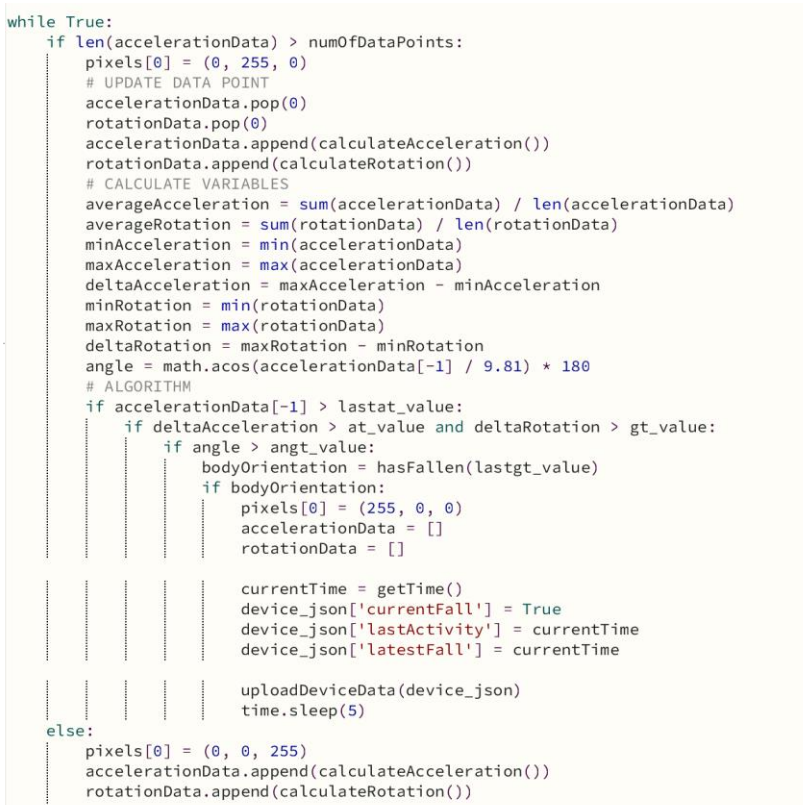
```
at_value = .1
gt_value = .1
angt_value = 45
lastat_value = 3
lastgt_value = 4
```

Figure 4. Values

"at_value" and "gt_value" are the thresholds for the delta between the highest and the lowest values of the last 100 linear and rotational acceleration data points. Essentially, a fall would register if there is a great enough change in both accelerations to surpass these values. "angt_value" is the angular threshold that the device must surpass, indicating the user's change in body orientation. "lastat_value" and "lastgt_value" are the linear and rotational acceleration thresholds for the last collected value of acceleration, to confirm whether the user is still in the process of accelerating due to a fall.

The algorithm is then produced with these values established. The following is a flowchart diagram of the algorithm (in the flowchart, "acceleration" is assumed to be linear):


Start
collect acceleration and rotational data
are 100 data points collected
No
Yes
replace oldest acceleration and rotation data with new data
calculate delta of acceleration and rotation of the min and max data points
No
No
No
is latest acceleration data greater than acceleration threshold?
Yes
do delta acceleration and delta rotation exceed their thresholds?
Yes
does orientation exceed threshold?
fall detected, upload to database
clear acceleration and rotation data



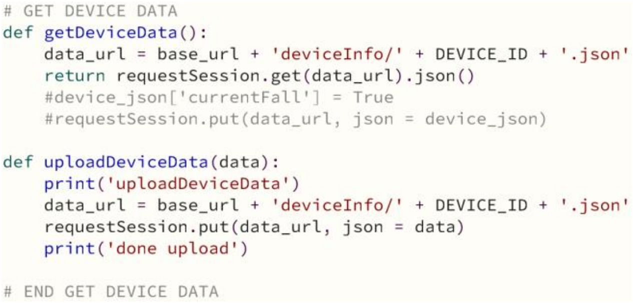
```
while True:
    if len(accelerationData) > numOfDataPoints:
        pixels[0] = (0, 255, 0)
        # UPDATE DATA POINT
        accelerationData.pop(0)
        rotationData.pop(0)
        accelerationData.append(calculateAcceleration())
        rotationData.append(calculateRotation())
        # CALCULATE VARIABLES
        averageAcceleration = sum(accelerationData) / len(accelerationData)
        averageRotation = sum(rotationData) / len(rotationData)
        minAcceleration = min(accelerationData)
        maxAcceleration = max(accelerationData)
        deltaAcceleration = maxAcceleration - minAcceleration
        minRotation = min(rotationData)
        maxRotation = max(rotationData)
        deltaRotation = maxRotation - minRotation
        angle = math.acos(accelerationData[-1] / 9.81) * 180
        # ALGORITHM
        if accelerationData[-1] > lastat_value:
            if deltaAcceleration > at_value and deltaRotation > gt_value:
                if angle > angt_value:
                    bodyOrientation = hasFallen(lastgt_value)
                    if bodyOrientation:
                        pixels[0] = (255, 0, 0)
                        accelerationData = []
                        rotationData = []

                        currentTime = getTime()
                        device_json['currentFall'] = True
                        device_json['lastActivity'] = currentTime
                        device_json['latestFall'] = currentTime

                        uploadDeviceData(device_json)
                        time.sleep(5)
    else:
        pixels[0] = (0, 0, 255)
        accelerationData.append(calculateAcceleration())
        rotationData.append(calculateRotation())
```

Figure 5. Flowchart diagram of the algorithm

First, the device collects a set of linear and rotational acceleration data points. Once the system has detected enough, it replaces the oldest acceleration data with a new data point. These data points are calculated from the "calculateAcceleration" and "calculateRotation" functions, which take the square root of the sum of the square of each dimension (ax2+ ay2+az2). From there, the algorithm takes the largest and smallest values in the new set and calculates the difference between them. Then, the system checks to see if the thresholds have been surpassed. It first determines whether the most recent linear acceleration exceeds the defined threshold. If so, it checks if the deltas of the linear and rotational accelerations are greater than their thresholds. If this is confirmed, the angle is checked by first determining whether the user has fallen greater than the angle threshold set, and then confirms that the most recent angular acceleration data point exceeds its threshold. At last, a fall is detected and uploaded to the database, and all data points are cleared while the device remains active.

Database and Communication

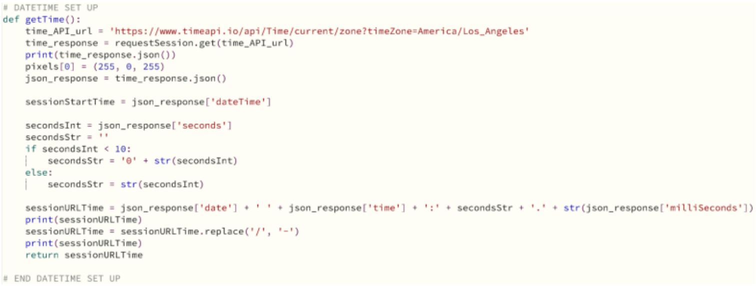
For exemplary purposes, Firebase is used to store user and device information, updated fall data, and fall histories. There are two main segments concerning the database: the manipulation of device data and timestamps.



```
# GET DEVICE DATA
def getDeviceData():
    data_url = base_url + 'deviceInfo/' + DEVICE_ID + '.json'
    return requestSession.get(data_url).json()
    #device_json['currentFall'] = True
    #requestSession.put(data_url, json = device_json)

def uploadDeviceData(data):
    print('uploadDeviceData')
    data_url = base_url + 'deviceInfo/' + DEVICE_ID + '.json'
    requestSession.put(data_url, json = data)
    print('done upload')

# END GET DEVICE DATA
```

Figure 6. Sending and receiving device data from database

In both receiving and uploading device data to the database, a URL is created [10]. Since a unique six-character (three letter three number) ID is assigned to each device and stored in the database, this ID is used to generate a unique link for the data of that specific device. With this URL labeled “data_url,” data can be received or updated with the respective functions shown above.

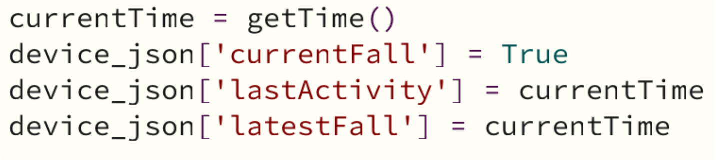


```
# DATETIME SET UP
def getTime():
    time_API_url = 'https://www.timeapi.io/api/Time/current/zone?timeZone=America/Los_Angeles'
    time_response = requestSession.get(time_API_url)
    print(time_response.json())
    pixels[0] = (255, 0, 255)
    json_response = time_response.json()

    sessionStartTime = json_response['dateTime']

    secondsInt = json_response['seconds']
    secondsStr = ''
    if secondsInt < 10:
        secondsStr = '0' + str(secondsInt)
    else:
        secondsStr = str(secondsInt)

    sessionURLTime = json_response['date'] + ' ' + json_response['time'] + ':' + secondsStr + '.' + str(json_response['milliSeconds'])
    print(sessionURLTime)
    sessionURLTime = sessionURLTime.replace('/', '-')
    print(sessionURLTime)
    return sessionURLTime

# END DATETIME SET UP
```

Figure 7. Structuring date and time to be uploaded to database

To track the time of falls, time is implemented via the timeapi.io live website. The time received must be reformatted before being uploaded to the database.

Device data and time are both used in the outcomes of the algorithm. For instance, when the system detects a fall, the “device_json” updates its values according to the fall event that occurred. The fall is set to true, and both the time of the last activity and latest fall are set to the present time. These values are ultimately read by the app, which updates according to the database.



```
currentTime = getTime()
device_json['currentFall'] = True
device_json['lastActivity'] = currentTime
device_json['latestFall'] = currentTime
```

Figure 8. Updating database data in fall occurrence

Mobile App

The mobile app is used as a live indication for when a fall has occurred [15]. The following are the key features of the app:

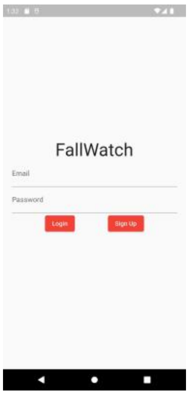
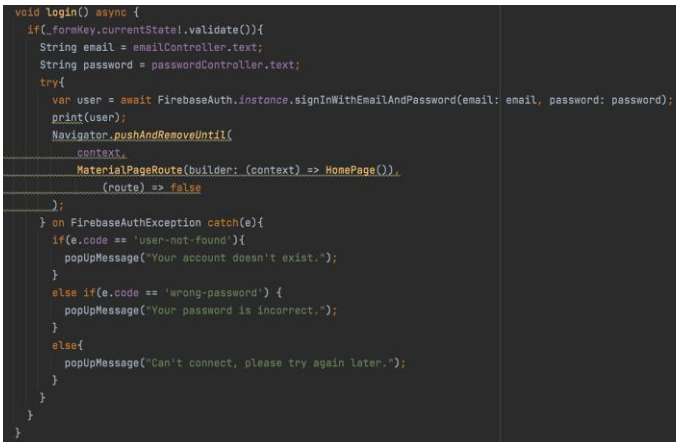


Figure 9. Login screen



```
void login() async {
  if(_formKey.currentState!.validate()){
    String email = emailController.text;
    String password = passwordController.text;
    try{
      var user = await FirebaseAuth.instance.signInWithEmailAndPassword(email: email, password: password);
      print(user);
      Navigator.pushAndRemoveUntil(
          context,
          MaterialPageRoute(builder: (context) => HomePage()),
              (route) => false
      );
    } on FirebaseAuthException catch(e){
      if(e.code == 'user-not-found'){
        popUpMessage("Your account doesn't exist.");
      }
      else if(e.code == 'wrong-password') {
        popUpMessage("Your password is incorrect.");
      }
      else{
        popUpMessage("Can't connect, please try again later.");
      }
    }
  }
}
```

Figure 10. Login screen code excerpt

The login interface is used to register the user in the database. Then, the fall device can be connected to the account via the unique six character device ID.

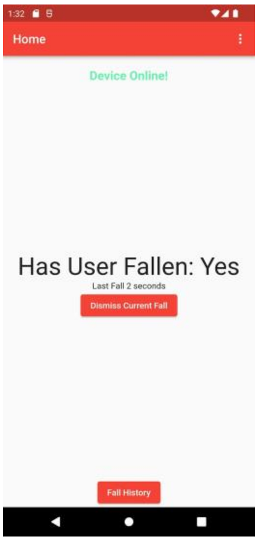
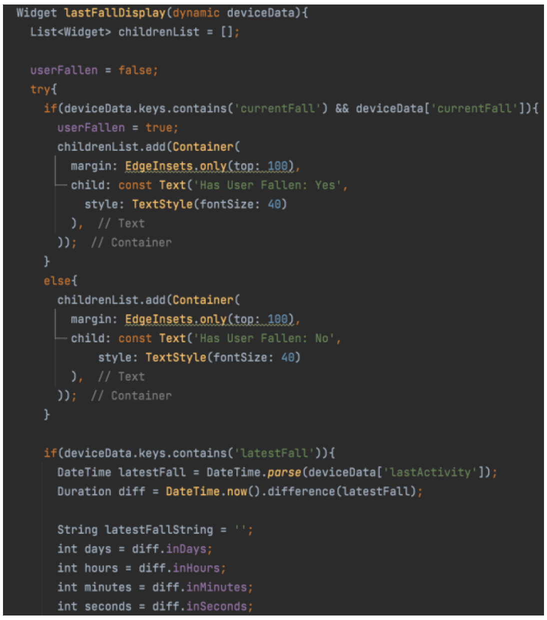


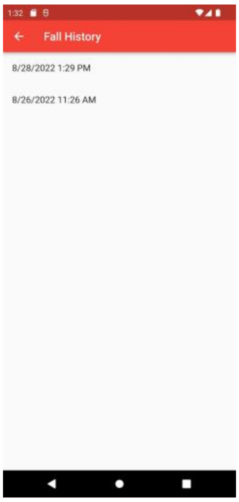
Figure 11. Homepage screen



```
Widget lastFallDisplay(dynamic deviceData){
  List<Widget> childrenList = [];

  userFallen = false;
  try{
    if(deviceData.keys.contains('currentFall') && deviceData['currentFall']){
      userFallen = true;
      childrenList.add(Container(
        margin: EdgeInsets.only(top: 100),
        child: const Text('Has User Fallen: Yes',
          style: TextStyle(fontSize: 40)
        ),  // Text
      ));  // Container
    }
    else{
      childrenList.add(Container(
        margin: EdgeInsets.only(top: 100),
        child: const Text('Has User Fallen: No',
            style: TextStyle(fontSize: 40)
        ),  // Text
      ));  // Container
    }

    if(deviceData.keys.contains('latestFall')){
      DateTime latestFall = DateTime.parse(deviceData['lastActivity']);
      Duration diff = DateTime.now().difference(latestFall);

      String latestFallString = '';
      int days = diff.inDays;
      int hours = diff.inHours;
      int minutes = diff.inMinutes;
      int seconds = diff.inSeconds;
```

Figure 12. Homepage code excerpt (detecting and displaying information about latest fall)

The homepage includes a few key components. At the top, the status of the device is displayed as whether the device is online and connected or offline. The center widget updates if the device detects a fall and shows the time elapsed since the last fall. The "Has User Fallen" status changes back to "No" after two minutes of no fall activity. There is also a button for the app user to dismiss the current fall, which manually reverts the status. At the bottom, there is a button for the fall history, which is as follows:

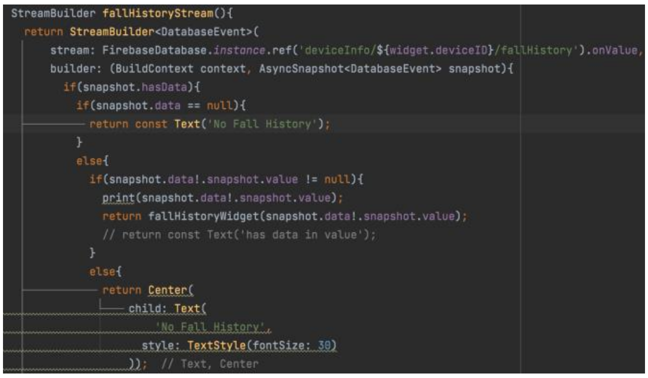


Fall History

8/28/2022 1:29 PM

8/26/2022 11:26 AM

Figure 13. Fall history screen



```
StreamBuilder fallHistoryStream(){
  return StreamBuilder<DatabaseEvent>(
      stream: FirebaseDatabase.instance.ref('deviceInfo/${widget.deviceID}/fallHistory').onValue,
      builder: (BuildContext context, AsyncSnapshot<DatabaseEvent> snapshot){
        if(snapshot.hasData){
          if(snapshot.data == null){
            return const Text('No Fall History');
          }
          else{
            if(snapshot.data!.snapshot.value != null){
              print(snapshot.data!.snapshot.value);
              return fallHistoryWidget(snapshot.data!.snapshot.value);
              // return const Text('has data in value');
            }
            else{
              return Center(
                  child: Text(
                    'No Fall History',
                  style: TextStyle(fontSize: 30)
                )); // Text, Center
```

Figure 14. Fall history screen code excerpt (retrieving data from database)

Here, a simple fall history log is displayed. The app gathers the device’s past falls and displays them in a list ordered from most recent downwards.

4. EXPERIMENT

- trials based on algo values, divided into 9 tests each (3 for each type of fall)
- data table and bar chart for success rate

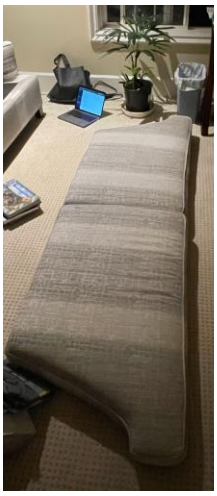
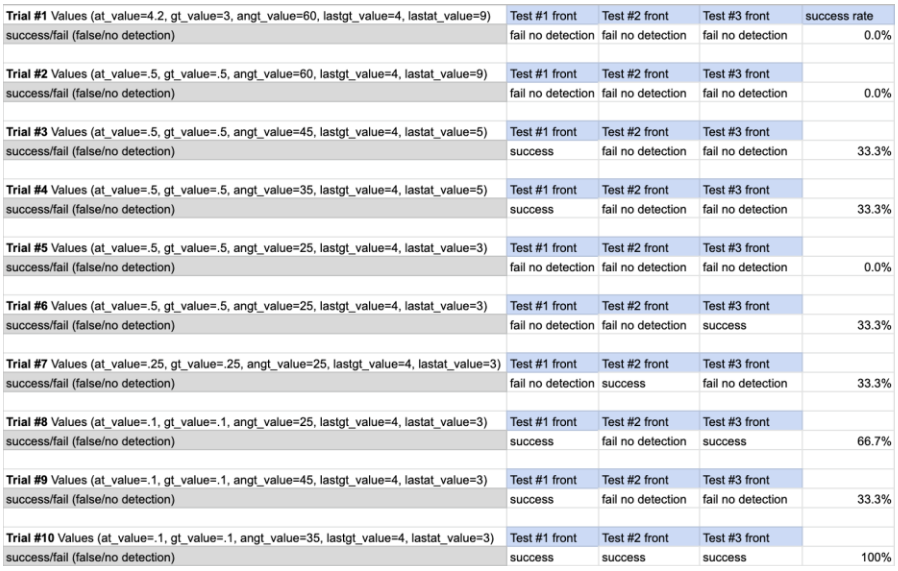


Figure 15. Fall area

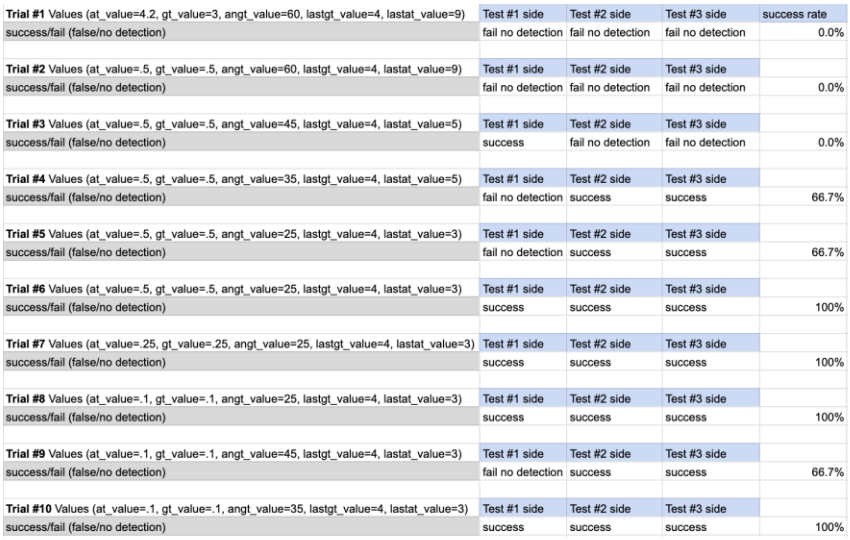
4.1. Experiment 1

To refine the algorithm, we tested falls until the set values produced accurate results. The device was plugged into a constant battery and taped on through the belt loop. An 8 foot fall landing area was set up and cushioned. Three mini-experiments were conducted in the setup, one for each type of fall (forward, sideways, backward). In each experiment, three falls were done for each set of threshold values, and the success of the device was recorded. Thresholds were then altered according to the sensitivity of the sensor for each subsequent set of tests. In total, 10 sets of values were tested.



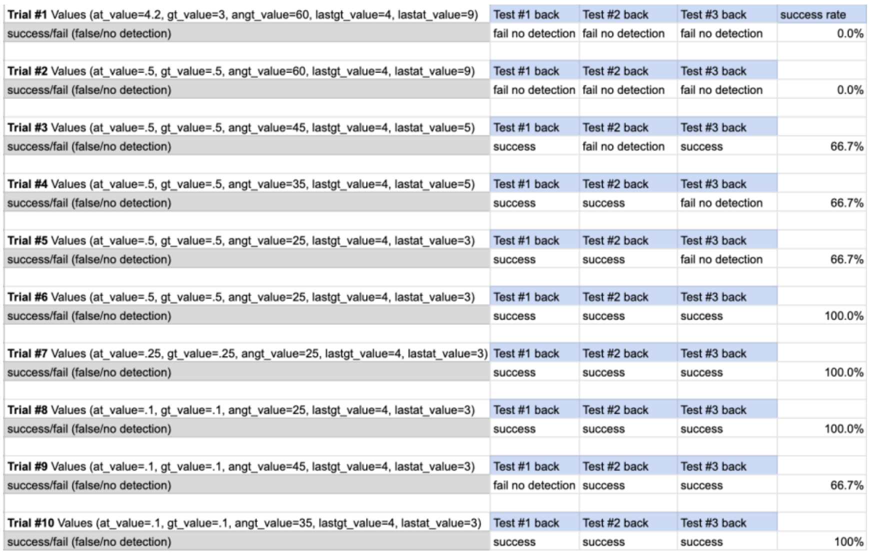
Trial #1 Values (at_value=4.2, gt_value=3, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 front	Test #2 front	Test #3 front	success rate
success/fail (false/no detection)	fail no detection	fail no detection	fail no detection	0.0%
Trial #2 Values (at_value=.5, gt_value=.5, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	fail no detection	fail no detection	fail no detection	0.0%
Trial #3 Values (at_value=.5, gt_value=.5, angt_value=45, lastgt_value=4, lastat_value=5)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	success	fail no detection	fail no detection	33.3%
Trial #4 Values (at_value=.5, gt_value=.5, angt_value=35, lastgt_value=4, lastat_value=5)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	success	fail no detection	fail no detection	33.3%
Trial #5 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	fail no detection	fail no detection	fail no detection	0.0%
Trial #6 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	fail no detection	fail no detection	success	33.3%
Trial #7 Values (at_value=.25, gt_value=.25, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	fail no detection	success	fail no detection	33.3%
Trial #8 Values (at_value=.1, gt_value=.1, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	success	fail no detection	success	66.7%
Trial #9 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	success	fail no detection	fail no detection	33.3%
Trial #10 Values (at_value=.1, gt_value=.1, angt_value=35, lastgt_value=4, lastat_value=3)	Test #1 front	Test #2 front	Test #3 front	
success/fail (false/no detection)	success	success	success	100%

Figure 16. Data for experiment #1 (forward fall)



Trial #1 Values (at_value=4.2, gt_value=3, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 side	Test #2 side	Test #3 side	success rate
success/fail (false/no detection)	fail no detection	fail no detection	fail no detection	0.0%
Trial #2 Values (at_value=.5, gt_value=.5, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	fail no detection	fail no detection	fail no detection	0.0%
Trial #3 Values (at_value=.5, gt_value=.5, angt_value=45, lastgt_value=4, lastat_value=5)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	success	fail no detection	fail no detection	0.0%
Trial #4 Values (at_value=.5, gt_value=.5, angt_value=35, lastgt_value=4, lastat_value=5)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	fail no detection	success	success	66.7%
Trial #5 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	fail no detection	success	success	66.7%
Trial #6 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	success	success	success	100%
Trial #7 Values (at_value=.25, gt_value=.25, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	success	success	success	100%
Trial #8 Values (at_value=.1, gt_value=.1, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	success	success	success	100%
Trial #9 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	fail no detection	success	success	66.7%
Trial #10 Values (at_value=.1, gt_value=.1, angt_value=35, lastgt_value=4, lastat_value=3)	Test #1 side	Test #2 side	Test #3 side	
success/fail (false/no detection)	success	success	success	100%

Figure 17. Data for experiment #2 (sideways fall)

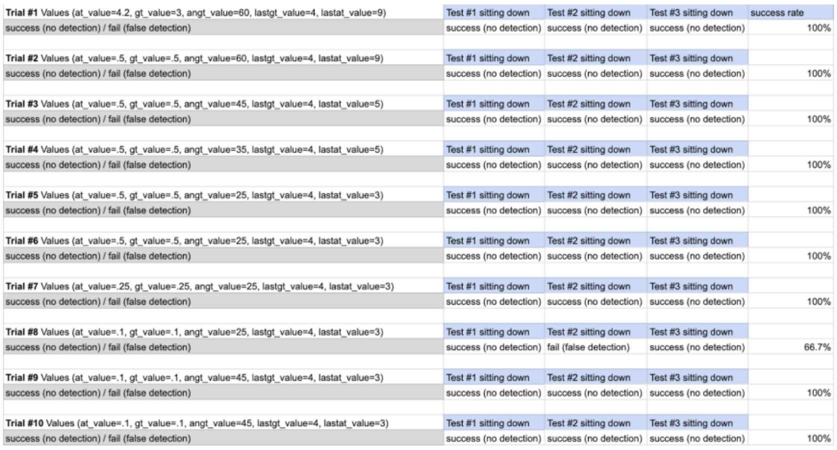


Trial #1 Values (at_value=4.2, gt_value=3, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 back	Test #2 back	Test #3 back	success rate
success/fail (false/no detection)	fail no detection	fail no detection	fail no detection	0.0%
Trial #2 Values (at_value=.5, gt_value=.5, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	fail no detection	fail no detection	fail no detection	0.0%
Trial #3 Values (at_value=.5, gt_value=.5, angt_value=45, lastgt_value=4, lastat_value=5)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	success	fail no detection	success	66.7%
Trial #4 Values (at_value=.5, gt_value=.5, angt_value=35, lastgt_value=4, lastat_value=5)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	success	success	fail no detection	66.7%
Trial #5 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	success	success	fail no detection	66.7%
Trial #6 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	success	success	success	100.0%
Trial #7 Values (at_value=.25, gt_value=.25, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	success	success	success	100.0%
Trial #8 Values (at_value=.1, gt_value=.1, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	success	success	success	100.0%
Trial #9 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	fail no detection	success	success	66.7%
Trial #10 Values (at_value=.1, gt_value=.1, angt_value=35, lastgt_value=4, lastat_value=3)	Test #1 back	Test #2 back	Test #3 back	
success/fail (false/no detection)	success	success	success	100%

Figure 18. Data for experiment #3 (backward fall)

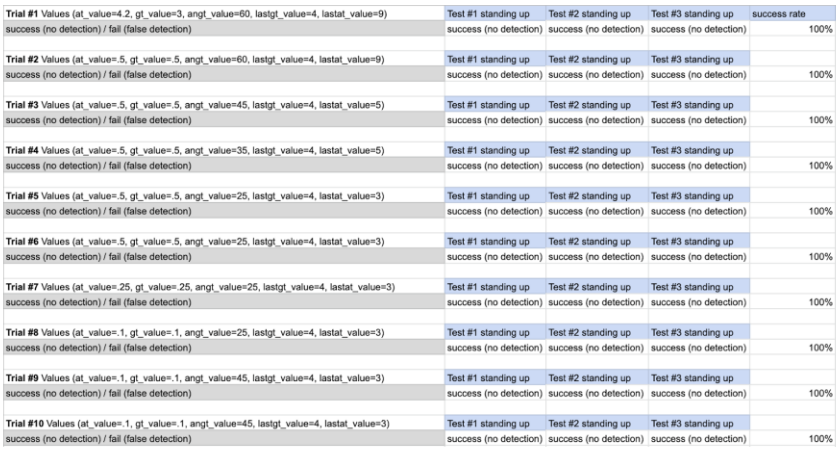
4.2. Experiment 2

In a second set of experiments, we needed to test for the event of false positives that mistake normal actions as falls. To do so, we tested each set of threshold values for three types of normal actions: standing up, sitting down, and walking. Each mini-experiment is structured the same way as the ones prior. The following are the trials and success rates for each action.



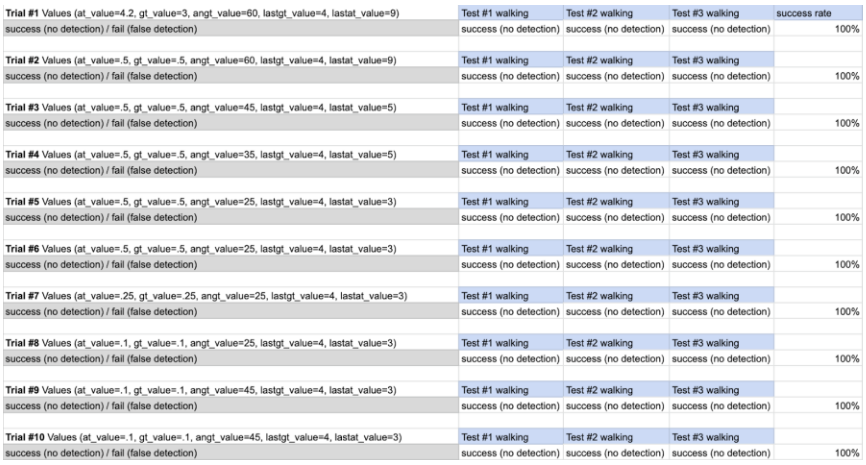
Trial #1 Values (at_value=4.2, gt_value=3, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	success rate
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #2 Values (at_value=.5, gt_value=.5, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #3 Values (at_value=.5, gt_value=.5, angt_value=45, lastgt_value=4, lastat_value=5)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #4 Values (at_value=.5, gt_value=.5, angt_value=35, lastgt_value=4, lastat_value=5)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #5 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #6 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #7 Values (at_value=.25, gt_value=.25, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #8 Values (at_value=.1, gt_value=.1, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	fail (false detection)	success (no detection)	66.7%
Trial #9 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #10 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 sitting down	Test #2 sitting down	Test #3 sitting down	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%

Figure 19. Data for experiment #4 (sitting down)



Trial #1 Values (at_value=4.2, gt_value=3, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 standing up	Test #2 standing up	Test #3 standing up	success rate
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #2 Values (at_value=.5, gt_value=.5, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #3 Values (at_value=.5, gt_value=.5, angt_value=45, lastgt_value=4, lastat_value=5)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #4 Values (at_value=.5, gt_value=.5, angt_value=35, lastgt_value=4, lastat_value=5)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #5 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #6 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #7 Values (at_value=.25, gt_value=.25, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #8 Values (at_value=.1, gt_value=.1, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #9 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #10 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 standing up	Test #2 standing up	Test #3 standing up	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%

Figure 20. Data for experiment #5 (standing up)



Trial #1 Values (at_value=4.2, gt_value=3, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 walking	Test #2 walking	Test #3 walking	success rate
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #2 Values (at_value=.5, gt_value=.5, angt_value=60, lastgt_value=4, lastat_value=9)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #3 Values (at_value=.5, gt_value=.5, angt_value=45, lastgt_value=4, lastat_value=5)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #4 Values (at_value=.5, gt_value=.5, angt_value=35, lastgt_value=4, lastat_value=5)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #5 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #6 Values (at_value=.5, gt_value=.5, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #7 Values (at_value=.25, gt_value=.25, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #8 Values (at_value=.1, gt_value=.1, angt_value=25, lastgt_value=4, lastat_value=3)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #9 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%
Trial #10 Values (at_value=.1, gt_value=.1, angt_value=45, lastgt_value=4, lastat_value=3)	Test #1 walking	Test #2 walking	Test #3 walking	
success (no detection) / fail (false detection)	success (no detection)	success (no detection)	success (no detection)	100%

Figure 21. Data for experiment #6 (walking)

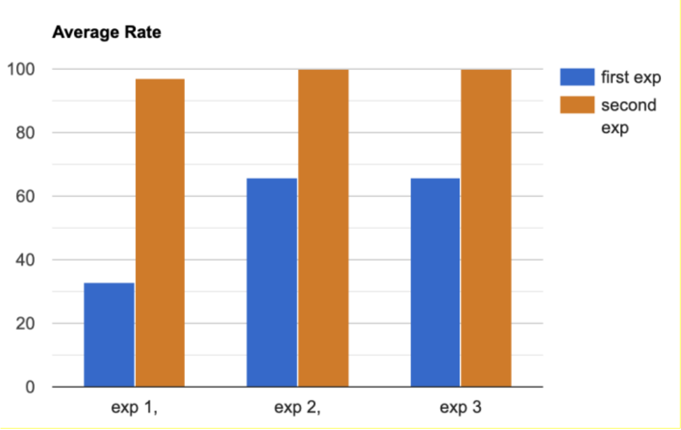


Figure 22. Average rate for 3 exp.

For our tests, the values of trial #10 (at_value = 0.1, gt_value = 0.1, angt_value = 35, lastgt_value = 4, lastat_value = 3) found the most success in both the fall and false test experiments. They were 100% accurate in the limited number of tests that we ran, and we acknowledge the low sample size as a potential pitfall. However, our results point out the fall detection capabilities of the tested sensors, which are meant to be affordable, compact, and noninvasive.

5. RELATED WORK

This work, published in 2009 by Zhou and his team, contains a similar approach [11]. Zhou's system features two TEMPO (Technology-Enabled Medical Precision Observation) 3.0 nodes placed on the front chest and thigh. Both have tri-axial gyroscope and accelerometer sensors with adequate sensitivities. The work defines and tests for four types of motion: activities of daily life (ADL), fall-like motions, flat falls, and falls onto inclined surfaces. Both the featured work and our algorithm use similar methods for calculating combined acceleration and body orientation. Both systems set the same types of thresholds for linear acceleration, rotational acceleration, and angle. However, the usage of two sensors can both be beneficial and detrimental. While it may produce more data for the algorithm to process, we must keep in mind that the positions of these sensors are in areas of the body (the chest and thigh) that are subject to high amounts of motion. Meanwhile, the hip stays relatively aligned with the body orientation, which may be more reliable to base a system off of.

Research done by this team utilizes a smartphone fitted in the chest shirt pocket as the tri-axial accelerometer and gyroscope [12]. Average linear and rotational acceleration are calculated in similar ways (square root of sum of dimensions squared), as well as angle. Both system algorithms also use similar methods of analysis. However, comparisons between the last instantaneous data point and the entire recorded data set are not made, which may cause more false positives with non-fall actions. Once again, the placement of the sensor may not be as reliable, but the data seems to show a high success rate nonetheless.

This research discusses energy consumption and battery life in different sensors that are used for fall detection purposes [13]. It proposes two types of sensor nodes: those that are custom made for fall detection and others that are general purpose devices. This is a unique angle when analyzing the integrating of fall devices. Since falls are spontaneous and can occur at any given moment, batteries need to last a sufficient length of time.

6. CONCLUSIONS

This system uses a tri-axis accelerometer and gyroscope attached to the hip to detect falls [14]. The sensors and controller communicate directly with a database, which stores previous activity and device information. A simple app was then created to simulate a common caretaker-elder relationship, where the app is updated and notified upon the device's trigger. This app-device environment serves as a proof of concept for an example of a fall response system.

To detect a fall, thresholds for linear acceleration, rotational acceleration, and angular change are set so that upon surpassing all three in the order of the algorithm, a fall is registered.

Because our work prioritized affordability as a key value, accuracy may suffer from the use of less expensive sensors (however the data collected does not indicate such an issue). Also, we have only created a post-fall detection algorithm like much research currently published [4]. A different method would likely need to be used to increase the speed at which the fall is initially detected. Finally, a different data sample would be beneficial to collect, one that incorporates the elder population as test subjects.

With a plan to implement automatic response capabilities to the device, developing a machine learning algorithm may be the optimal approach to creating not only a fall detection system, but a fall prevention product. Although implementing an airbag deployment system seems intuitive (and has been experimented with), this design has not been readily available on the market due to regulation constraints. In the expansive market of fall injuries, innovative steps must be made to ensure the safety of countless lives in the future.

REFERENCES

[1] Hobbs, Frank B. "The elderly population." US Census Bureau Population Profile of the United States, http://www. census. gov/population/www/pop-profile/elderpop. html (2001).

[2] Ellis, J. Nigel. "What is Fall Protection?." ASSE Professional Development Conference and Exposition. OnePetro, 2000.

[3] Mubashir, Muhammad, Ling Shao, and Luke Seed. "A survey on fall detection: Principles and approaches." Neurocomputing 100 (2013): 144-152.

[4] Igual, Raul, Carlos Medrano, and Inmaculada Plaza. "Challenges, issues and trends in fall detection systems." Biomedical engineering online 12.1 (2013): 1-24.

[5] Wang, Xueyi, Joshua Ellul, and George Azzopardi. "Elderly fall detection systems: A literature survey." Frontiers in Robotics and AI 7 (2020): 71.

[6] Southern, W. Thomas, and Eric D. Jones. "Types of acceleration: Dimensions and issues." A nation deceived: How schools hold back America's brightest students 2 (2004): 5-12.

[7] Negahban, Arash, and Chih-Hung Chung. "Discovering determinants of users perception of mobile device functionality fit." Computers in Human Behavior 35 (2014): 75-84.

[8] Wedekind, Hartmut, and George Zoerntlein. "Prefetching in realtime database applications." ACM SIGMOD Record 15.2 (1986): 215-226.

[9] Pan, Jiapu, and Willis J. Tompkins. "A real-time QRS detection algorithm." IEEE transactions on biomedical engineering 3 (1985): 230-236.

[10] Berners-Lee, Tim, Larry Masinter, and Mark McCahill. Uniform resource locators (URL). No. rfc1738. 1994.

[11] Li, Qiang, et al. "Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information." 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks. IEEE, 2009.

[12] Rakhman, Arkham Zahri, and Lukito Edi Nugroho. "Fall detection system using accelerometer and gyroscope based on smartphone." 2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering. IEEE, 2014.

[13] Gia, Tuan Nguyen, et al. "IoT-based fall detection system with energy efficient sensor nodes." 2016 IEEE Nordic Circuits and Systems Conference (NORCAS). IEEE, 2016.

[14] ㅔ에. "Physical activity recognition using a single tri-axis accelerometer." Proceedings of the world congress on engineering and computer science. Vol. 1. 2009.

[15] Joorabchi, Mona Erfani, Ali Mesbah, and Philippe Kruchten. "Real challenges in mobile app development." 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. IEEE, 2013.

AUTHOR INDEX