CSCI 2022 BOOK of ABSTRACTS

The 2022 World Congress in Computer Science, Computer Engineering, and Applied Computing
CSCE 2022

https://american-cse.org/csce2022/
July 25-28, 2022
Luxor Hotel (MGM Property), 3900 Las Vegas Blvd. South, Las Vegas, 89109, USA

Table of Contents

Note that the titles of papers, the authors' names and the abstracts that appear in this book ("Book of Abstracts") were extracted from the papers that were submitted to the EVALUATION web portal (i.e., extracted from the first draft submissions). The official published proceedings/book will have any and all changes/revisions that authors may have done.

KEYNOTE ADDRESSES

(The Keynote lectures are open to all participants)

CONGRESS WELCOME REMARKS

Professor Emeritus Hamid R. Arabnia
(Chair, Steering Committee & Coordinator), School of Computing, University of Georgia, USA;
Editor-in-Chief, The Journal of Supercomputing (Springer);
Fellow, Center of Excellence in Terrorism, Resilience, Intelligence & Organized Crime Research (CENTRIC)

Abstract: N/A

ARTIFICIAL EMOTIONS FOR AI AVATARS: ENHANCING HUMAN/AI INTERACTIONS AND COLLABORATION

Dr. James A. Crowder
Systems Fellow AT&E,
CAES Advanced Technology and Engineering (AT&E), USA
(Formerly at Raytheon Technologies. Recipient of technical achievement awards from Lockheed Martin, General Dynamics, Northrup Grumman, & Raytheon Technologies)

Abstract: On-line avatars that claim to be based on artificial intelligence are prevalent across the internet, including chatbots and applications claiming to be on-line AI-driven counselors. The majority of these avatars are script-driven applications that utilize word/phrase searches to determine scripted responses and inquiries (questions) they pose back to the human. Attempts to create fully learning-driven chatbots, like the Microsoft's Twitter Chatbot Tay (2016) that was turned into a racist entity in around 24 hours, fail at effective human/AI communications. Most chatbots are basically the intelligence and mentality of around a 6-year-old child. They can do basic tasks but are easily confused. As we move closer and closer to real cognitive AI entities with near human intelligence (we will use the term synthetic intelligence) one issue that will be of extreme importance for synthetically intelligent entities is the ability to understand, process, and respond in kind to human emotions. Without an understanding and processing of basic human emotions, chatbots, AI-avatars, and synthetically intelligent entities will always fail at human/AI interactions, since emotions are always at the heart of human communication, learning, and reasoning. We will explore constructs and the mathematics of synthetic emotions that will allow synthetically intelligent entities to learn, process, and respond to humans in ways that demonstrate the artificial emotional intelligence necessary for effective human/AI communications.

SYSTEM & METHODS FOR OPTIMAL SYNCHRONIZED WORKFLOW BASED COMMUNICATION & A FRAMEWORK FOR SEAMLESS INTEGRATION OF AI WITH WORKFLOW

Massoud Alibakhsh
Founder & CEO, Xeba Technologies, Atlanta Georgia, USA
www.xeba.tech

Abstract: Email and Social Channel models of communication promised to revolutionize and simplify communication at work. Instead, they have created a sense of enslavement and overwhelmed us with noisy and unrelated distractions. This talk introduces a new and elegant paradigm for human communication at work. One that guarantees optimality in that you will never receive a message that isn't relevant to you. It also presents a framework for a harmonious merging of AI with people in the workplace to aid us in every aspect of our work in pleasantly surprising ways!

HIGH TECH NOSTALGIA TO THE FUTURE OF COMPUTING

Prof. Douglas D. Hodson
Professor of Computer Engineering,
Air Force Institute of Technology (AFIT), Wright-Patterson AFB, Ohio, USA

Abstract: A somewhat nostalgic retrospective on growing up in the PC era, how technology has changed, and what the future might hold. By recognizing how current configurable hardware technologies provide the capability to emulate past computer hardware designs, and how current high level synthesis tools provide a capability to implement software as hardware, one can envision a future computer that runs programs that configures the hardware as needed.

HIGH-LEVEL DEVELOPMENT OF DISTRIBUTED REAL-TIME APPLICATIONS USING MOBILE CLOUD & SOFTWARE-DEFINED NETWORKING

Prof. Dr. habil. Sergei Gorlatch
Professor of Computer Science, University of Muenster, Germany

Abstract: We address the software development process for challenging class of Real-Time Online Interactive Applications (ROIA). ROIA are distributed applications connecting a potentially very high number of users who interact with the application and with each other in real time, i.e., a response to a user's action happens virtually immediately. Typical representatives of ROIA are massively multiplayer online computer games, advanced simulation-based e-learning and serious gaming. These applications are characterized by high performance and Quality-of-Service (QoS) requirements, such as: short response times to user inputs (about 0.1-1.5 sec); frequent state updates (up to 100 Hz); large and frequently changing numbers of users in a single application instance (up to tens of thousands simultaneous users). This talk will present our high-level development framework RTF and address two challenging issues in future Internet-based ROIA applications: a) using Mobile Cloud Computing for allowing high application performance when a ROIA application is accessed from multiple mobile devices, and b) managing dynamic QoS requirements of ROIA applications by employing the emerging technology of Software-Defined Networking (SDN).

AUTOMOTIVE INTELLIGENCE FOR CONNECTED, SHARED AND AUTONOMOUS MOBILITY OF THE FUTURE

Prof. Dr. Eng. George Dimitrakopoulos
Harokopio University of Athens, Greece;
Infineon Technologies AG, Munich, Germany

Abstract: Mobility and transport continue to attract immense research and development interest. Despite the innumerable achievements, there are still lots of challenges to be overcome. In this context, the keynote is about the future of mobility in relation to 4 major trends in mobility, namely the ECAS 2030 vision (Electric, Connected, Autonomous and Shared). It will cover:

a) Some generic technical and user requirements and specifications for ECAS vehicles
b) hardware developments, e.g. accelerators for connected and shared mobility
c) management algorithms enabled by AI tools and methods for decision making supporting in-vehicle intelligence
d) how we can support penetration of new relevant technologies that will be accepted from the society
e) the progress beyond the state of the art reflected in current European research and development funded projects.

Overall, the presentation aims at providing an overview of research achievements and challenges in the area of intelligent transport systems and their continuous transformation towards systems for connected and shared mobility.

DEEP LEARNING COMPUTER VISION TO ADVANCE TOTAL JOINT ARTHROPLASTY (TJA) RESEARCH

Dr. Ahmad P. Tafti
Director, Health + Explainable AI (HexAI) Research Laboratory (https://hexailab.github.io/)
Graduate Program Director; University of Southern Maine, Portland, USA

Abstract: Orthopedic surgical procedures, and particularly total knee/hip arthroplasty (TKA/THA), are the most common and fastest growing surgeries in the United States. Almost 1.3 million TJA procedures occur on a yearly basis and more than 7 million Americans are currently living with artificial knee and/or hip joints. The widespread adoption of x-ray radiography and their availability at low cost, make them the principal method in assessing TJA and subtle TJA complications, such as osteolysis, implant loosening or infection over time, enabling surgeons to rule out complications and possible needs for revision surgeries. Rapid yet, with the growing number of TJA patients, the routine clinical and radiograph follow-up remain a daunting task for most orthopedic centers. It becomes an overwhelming amount of work, on a human scale, when we consider a radiologist or surgeon presented with the vast number of medical images daily. Smart computational strategies, such as artificial intelligence and deep learning computational vision are thus required to analyze arthroplasty radiographs automatically and objectively, enabling both naive and experienced practitioners to perform radiographic follow-up with greater ease and speed. In this talk, we will be discussing the effectiveness of modern computer vision methods to advance TJA research. We, together, will explore what computational vision components do in TJA research and how.

The 6th International Conference on Applied Cognitive Computing (ACC'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-ACC
https://american-cse.org/csce2022/

Incorporating Topological Complexity into a Multilayer Perception

Kenneth Brezinski, Ken Ferens
Department of Electrical and Computer Engineering, University of Manitoba, Canada

Abstract: Our paradigm for incorporating complexity into model training is done through computing the topological complexity of the intermediate pre- and post activation layers in a multilayer perceptron architecture. The key to our implementation is the partitioning of the embeddings at each layer, followed by computing the standard deviations of the training examples from a static and dynamic centroid, and finally, solving for the power law relationship between the number of training example considered at each scale and the cumulative standard deviation. The proposed complexity measure was tested on imbalanced datasets (1% and 5% positive), on pre- and post-activation outputs, and with different order of the norm for the distance metric. The procedure demonstrated higher validation accuracy at the cost of increased validation loss, which was largely attributed to exploding activations and gradients which were not normalized or participants in backprogation. We propose our complexity layer which will resolve these issues in theory using learnable parameters.

The Rasmussen Cognition Model as a Framework for a Modified OODA Loop

James Crowder, Michael Hirsch
CAES Advanced Technology and Engineering, Colorado Springs, USA;
ISEA TEK LLC, Maitland, Florida, USA

Abstract: Cognitive architectures are beginning to play an increasingly important role as mechanisms for facilitating real-time, continuous learning in intelligent systems. What is presented here is the architecture for a cognitive framework for applications like reconnaissance and surveillance. Here we borrow from the military concept of operations management and start with a modification of the DoD Observe, Orient, Decide and Act (OODA) loop. We add a machine learning component and adapt this for processing and execution of countermeasures to sensed threats within a Cognitive, Autonomous Processing System (CAPS). Our concept, the Observe, Orient, Decide, Act, and Learn (OODAL) loop makes use of the Rasmussen Cognition Model (RCM) to determine, during the observe and orient phases, whether the situation constitutes knowledge-based, skills-based, or rules-based decisions, which will affect the possible actions available to the system, whether autonomous or semi-autonomous. We utilize a reinforcement learning component in the system that allows the AI entity to learn from experience (skills-based learning) whether its decisions were of benefit or problematic.

A Novel Solution of Enhance Extreme Learning Machine in Temperature Prediction of a Building

Ramanan Selvaratnam, Nawzat Sadiq Ahmed, Abeer Alsadoon, Razwan Mohmed Salah, P.W.C. Prasad, Oday D. Jerew
School of Computing and Mathematics, Charles Sturt University (CSU), Wagga Wagga, Australia;
Department of Information Technology Management, Duhok Polytechnic University - Duhok, KRG, Iraq;
School of Computer Data and Mathematical Sciences, University of Western Sydney (UWS), Sydney, Australia;
College of Science, Department of Computer Science, The University of Duhok, Duhok, KRG, Iraq;
Asia Pacific International College (APIC), Information Technology Department, Sydney, Australia;
Kent Institute Australia, Information Technology Department, Sydney, Australia

Abstract: Machine learning (ML) techniques have been used to reduce energy utilization of indoor temperature regulation. However, one of the main limitations of current implementations is the poor prediction accuracy and performance. Therefore, and this research proposes a system by enhancing Extreme Learning Machine (ELM) to reduce unnecessary energy consumed and predict indoor temperature using the Generative Adversarial Network. This study focuses on the reduction of noise impact on the prediction accuracy using GANs. Historical weather dataset used in our simulation studies National Oceanic and Atmospheric Administration collected the dataset. The simulation results show that The prediction accuracy increased to 83% from 79%, and the difference in total processing time between the proposed solution and the state-of-the-art approach was between 2 and 21 milliseconds (ms). The proposed solution can predict future indoor temperature against time series and analyze multiple buildings and weather data.

Risk-Based Cognitive Decision-Making Functionality for the Reconfiguration of the Level of Autonomy in Highly Automated Vehicles

Ilias Panagiotopoulos, Konstantina Karathanasopoulou, George Dimitrakopoulos
Department of Telematics and Informatics, Harokopio University of Athens (HUA), Athens, Greece

Abstract: Highly Automated Vehicles (HAVs) are expected to fundamentally change transportation by increasing traffic flow efficiency, comfort and driving productivity. HAVs dispose decision-making functionalities that undertake the responsibility to keep them safe. With the help of recent developments in cognitive management techniques, knowledge based decision making can further increase the velocity and efficiency of decisions. This paper introduces the factor of risk associated with every possible decision, by incorporating in the decision making process the "a priori risk assessment" and enriching the decisions upon the most appropriate Level of Autonomy (LoA). The cognitive nature of the proposed analysis is based on previous knowledge by implementing Naive Bayes principles. The effectiveness of the proposed methodology is showcased through simulations.

Classifying SARS-CoV-2 and Common Co-infections from Genome Assemblies

Mohaimen Rahman, Kenneth Brezinski, Ken Ferens
Department of of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada

Abstract: Evidence has suggested that co-infections among Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) patients lead to worsening health outcomes. Polymerase chain reaction (PCR) testing has widespread use in the systematic identification of SARS-CoV-2 strains; however, it is unable to alert to the presence of multiple coinfections simultaneously without specialized equipment. One solution is genome sequences which act as a unique fingerprint of base pairs that can provide a form of quick characterization. To this end, we investigated a k-mer approach to classify genome sequences of SARS-CoV-2 and its common coinfections, as well as a Human genome sequence. We aim to provide a simplified classification approach that balances validation time while limiting hyperparameter tuning. Our approach achieved F1 scores in excess of 0.97, and perfect scores between the common co-infections. We demonstrated a simple 5-base sub-sequencing scheme has the power to differentiate over 7.91 million sequences from almost 20 thousand genome assemblies.

Classification Method and its Effective Use for Web Advertisements Based on Pupil Fluctuations

Ryuta Yamashita, Tadaaki Shimizu, Natsuki Yoshinaka
Graduate School of Engineering, Tottori University, Tottori, Japan;
Cross-informatics Research Center, Tottori University, Tottori, Japan;
Department of Engineering, Graduate School of Sustainability Science, Tottori University, Tottori, Japan

Abstract: In the "Awareness Survey on Internet Advertising," it was reported that nearly 80% of web users feel that web advertising is an obstacle to their browsing experience. Current recommendation systems are based on the user ID and website brows- ing history, among other factors. However, it is difficult to provide appropriate advertisements to users using only such information. In this paper, we analyze the results of web manipulation experiments by subject and classified the features of pupillary variability using a self-organizing map and showed that these features are common to many subjects. In addition, we considered how to present advertisements that could be appropriately appealing to users based on the results.

DGAN: A Novel Dynamic Graph Attention Networks for Accurately Traffic Prediction

Renyi Chen, Huaxiong Yao
Central China Normal University, Hubei, P. R. China

Abstract: Accurate prediction of future road traffic conditions is great significance for pedestrian travel and social development. Due to the complex spatial-temporal dependencies of the traffic network, there exist a huge challenge to model. In this paper, we propose a novel dynamic graph attention network (DGAN) for accurate traffic prediction to address the above mentioned challenges. Complex spatial-temporal information is extracted by stacking multi-layer of polymorphic spatial-temporal blocks, which include adaptive graph convolution network and spatial-temporal attention block. In addition, we introduce a gate function to control the flow of residual information. This gating mechanism helps determine how much previous residual information affects the next time segments. Extensive experiments on two real-world road network traffic datasets show that the DGAN achieves the state-of-the-art results.

Creating the Dictionary of Abstract Nouns: A Comparison with Machine Ranking Data

Yulia A. Volskaya
Kazan Federal University, Kazan, Russian Federation

Abstract: Concreteness and abstractness are important concepts for many studies that go beyond traditional linguistics. Abstract words are studied in the context of neurolinguistic, psycholinguistic, cognitive and other studies. In this regard, dictionaries are created for different languages, in which for each input unit the degree of abstractness or concreteness attribute is indicated. Such dictionaries have already been created for English, Dutch, and are also being created for Croatian, Italian, Chinese and other languages. Word rankings are usually collected through surveys among native speakers, but machine rankings of abstractness are also created in parallel. This paper describes the methodology of creating a dictionary of abstract units for the Russian language by conducting surveys. The respondents\' scores are compared with three versions of machine ratings of Russian language abstractness. The correlation coefficient between the human ratings and the machine ratings is considered; the words with the largest difference between the machine and human ratings are analyzed.

The 23rd International Conference on Bioinformatics & Computational Biology (BIOCOMP'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-BIOCOMP
https://american-cse.org/csce2022/

Temperature Effect on the Structural Dynamics of SARS-CoV-2 Nucleocapsid Domain

Samuel M. Naandam, Gideon K. Gogovi, Samuel Kakraba
Department of Mathematics, University of Cape Coast, Cape Coast, Ghana;
Department of Mathematics and Statistics, University of Houston - Downtown, Houston, Texas, USA;
Department of Mathematics and Statistics, Eastern Kentucky University, Richmond, Kentucky, USA

Abstract: The coronavirus machinery is complex and may have other stabilizing mechanisms and protein-protein or protein-solvent interactions to cope with environmental difficulty that is brought forth during the transition between summer, fall, winter, spring and back to summer. In this paper, we present molecular dynamics investigation on the effect of temperature on the structural dynamics and energetics of SARS-CoV-2 nucleocapsid protein. This study is aimed at understanding and explaining the sensitivity of this domain to different temperature environments. In general, we found that stability of the protein increases with increasing temperature. The trend, however, was different for the 293.15K system. Structural characterization of the atomic trajectories however did not offer a clear understanding of the structural and dynamical behavior of the domain under different temperature conditions studied along time. Evidence of high individual residue flexibility was also observed in a region made up of a pocket of aromatic, hydrophobic amino acids. Our results suggest that this SARS-CoV-2 protein is well adapted to habitable temperatures and exhibit extreme stability.

Practical Screening Method of Cancer Gene Diagnosis - Four Universal Data Structures of Discriminant Data

Shuichi Shinmura
Professor Emeritus Seikei University, Sakasai Kashiwa City, Chiba, Japan

Abstract: In 2015, we completed the discriminant theory proper for medical diagnosis (Theory1). Revised IP Optimal LDF (RIP) discriminated six first-generation arrays (old arrays). The minimum numbers of misclassifications (MNMs) are zero and linearly separable data (LSD). Our Program3 and Program4 coded in LINGO can split the array into many Small Matryoshkas (SMs), and Basic Gene Sets (BGSs). SMs and BGSs are multivariate oncogenes to discriminate between normal and cancer patients (or two cancer classes) with a few genes. We confirm these facts by analyzing 163 additional new arrays after 2007. We find four universal discriminant data structures (Fact3) of cancer gene data analysis (Theory2). We expect physicians use our Theory2 as the screening method before cancer diagnosis. However, it is difficult for physicians to understand Theory2 in detail. Therefore, we develop a practical screening method and analyze six new arrays among 163 new microarrays. This paper analyzes six old arrays by our screening method and evaluates the quality of both new and old arrays.

Practical Screening Method for Cancer Gene Diagnosis - How to Choose Cancer and Normal Patients by Four Principles

Shuichi Shinmura
Professor Emeritus Seikei University, Sakasai Kashiwa City, Chiba, Japan

Abstract: We developed a new theory of discriminant analysis (Theory1). Physicians can use it for practical medical diagnoses. Only Revised IP Optimal-LDF (RIP) obtains the minimum number of misclassification (MNM) and Error Rate (ER). RIP can discriminate linearly separable data (LSD) theoretically. It discriminated against 169 microarrays with two classes and found that 169 MNMs are zero and LSD. It can split high-dimensional arrays into many small LSD with less than n (patient's number) genes that are the candidates of multivariate oncogenes. We completed a new theory of high-dimensional gene data analysis (Theory2). A 100-fold or 10-fold Cross-Validation (Method1) can rank all candidates for the importance of diagnosis. Thus, if physicians firstly use Theory2 as the screening method, they can start their medical studies with the correct small sizes of candidates. This paper analyzes four arrays in detail and proposes correctly choosing cancer and normal patients using four principles.

Mathematical Dynamics of Three Kinds of Genetic Code Equivalences

Matthew He
Department of Mathematics, Halmos College of Arts and Sciences, Nova Southeastern University, Florida, USA

Abstract: In biology, a genetic system provides the self-reproduction of biological organisms in their generations. In mathematics, the "golden ratio" (or the "divine proportion") and its properties were a mathematical symbol of self-reproduction. The biological system divides the genetic four-letter alphabet (A, C, G, T/U) into various three pairs of letters. There are three kinds of genetic equivalences among these four-letter alphabets (1. A=C, G=U; 2. A=G, C=U; 3. A=U, C=G). In this paper, we investigate the dynamics of these three kinds of genetic code equivalences. We show that each equivalence has its own mathematical dynamics by introducing three 2x2 matrices associated with three genetic code equivalences. These dynamics include attracting fixed points, repelling fixed points, basin of attractions, Julia sets and corresponding Mandelbrot sets. We further study three (8x8) matrices associated with three special ratios (unity ratio, ring ratio, and golden ratio) and three genetic equivalences. We show that these matrices associated with these special ratios are symmetric and doubly stochastic matrices.

Self-supervised Multimodal Pre-training for Lung Adenocarcinoma overall Survival Prediction

Francisco Carrillo-Perez, Marija Pizurica, Ignacrio Rojas, Kathleen Marchal, Luis Javier Herrera, Olivier Gevaert
Department of Computer Architecture and Technology,
University of Granada. C.I.T.I.C., Periodista Rafael Gomez Montero, Granada, Spain;
Stanford Center for Biomedical Informatics Research (BMIR), Department of Medicine,
Stanford University, Stanford, California, USA;
Internet Technology and Data science Lab (IDLab), Ghent University, Technologiepark-Zwijnaarde, Gent, Belgium

Abstract: The collection of multiple modalities of cancer data has increased over the years, allowing research in complex problems such as cancer prognosis. However, given the high-dimensionality of biological data, efficiently training machine learning models when scarce samples are available is still challenging. In this work we propose a novel multimodal self-supervised learning framework based on neural networks for survival analysis and we evaluate it in a few-shot learning setting for lung adenocarcinoma prognosis. We show that the multimodal self-supervised pre-training is more effective than regular pre-training or training from scratch for two modalities (RNA-Seq and Whole Slide Imaging) when few samples are available. With the multimodal self-supervised learning framework, the relation between the modalities is learned in a pretext task and the leveraged information is successfully used for the relevant downstream task for both modalities, showing the potential of the proposed methodology.

A Modification Of Chatterjee's Correlation Test and its Applications to Transcriptomics Analysis

Qingyang Zhang
Department of Mathematical Sciences, University of Arkansas, Fayetteville, Arkansas, USA

Abstract: Chatterjee (2021) introduced a new correlation test that has attracted much attention over the past year. The test is distribution-free, asymptotically normal under independence, and consistent against all fixed alternatives, facilitating its applications to large-scale association studies. However, as pointed out by several recent studies, Chatterjee\'s test is unfortunately sub-optimal and they call for variants. We answer this call by proposing a modification of Chatterjee\'s test statistic based on weighted moving averages. Numerical studies show that our test consistently outperforms Chatterjee\'s test and a variant proposed by Lin and Han (2021). We illustrate the new method on two popular transcriptomics datasets including the yeast cell cycle data and the TCGA ovarian cancer data.

Efficient Prediction of Tumor Gene Association using Machine Learning and Statistical Methods

Surabhi Mishra, Abhishek Garain, Mahua Bhattacharya
Department of IT, ABV-Indian Institute of Information Technology and Management, Gwalior, India

Abstract: The tumor-gene associations have high importance since they can help us discover the mechanisms of particular cancer. However, it is exhausting to find a connection between a gene and a tumor experimentally. The proposed scheme drives the tumor-gene associations for the DNA methylation dataset through gene-gene interaction using the heat diffusion algorithm. The network's active nodes that participate for the above process considers as seed nodes, and they select using the type 2 fuzzy model. This scheme tries to determine and visualize the connection of cancer from an ancestor to offspring. For the BioGRID database, this algorithm predicts the links between tumor samples and genes with a mean AUC-ROC score of 0.84 and a high AUC-ROC score of 0.90.

Embedding Of Functional Human Brain Networks on a Sphere

Moo K. Chung, Zijian Chen
University of Wisconsin-Madison, Wisconsin, USA

Abstract: Human brain activity is often measured using the blood-oxygen-level dependent (BOLD) signals obtained through functional magnetic resonance imaging (fMRI). The strength of connectivity between brain regions is then measured as a Pearson correlation matrix. As the number of brain regions increases, the dimension of matrix increases. It becomes extremely cumbersome to even visualize and quantify such weighted complete networks. To remedy the problem, we propose to embed brain networks onto a sphere, which is a Riemannian manifold with constant positive curvature. The Matlab code for the spherical embedding is given in https://github.com/laplcebeltrami/sphericalMDS.

How Information Theory shall be Generalized for Neural Communication

Janos Vegh, Adam Jozef Berki
Kalimanos BT, Debrecen, Hungary;
Department of Neurology, Semmelweis University, Budapest, Hungary

Abstract: Neuroscience extensively uses information theory to describe neural communication, among others, to calculate the amount of information transferred in neural communication and to attempt cracking of its coding. However, there are severe debates on how information is represented in the brain and during transmission inside the brain. The neural information theory attempts to use the assumptions of electronic communication despite the experimental evidence that the neural spikes carry information on non-discrete states, they have shallow communication speed, and the spikes\' timing precision matters. Furthermore, in biology, the communication channel is active, which enforces an additional neural bandwidth limitation to the neural information transfer. The paper revises the notions needed to describe information transfer to discuss technical and biological communication systems. It demonstrates that biology uses Shannon\'s idea outside of its range of validity and introduces an adequate interpretation of information. The generalized theory describes both kinds of communication, and the classic information theory is the particular case of the generalized information theory.

Two-level Classification of Mental Stress Using DenseNet

Yu-Hung Tsai, Sheng-Kuang Wu, Shyr-Shen Yu, Meng-Hsiun Tsai
Department of Computer Science & Engineering, National Chung Hsing University, Taichung City, Taiwan;
Department of Sport Performance, National Taiwan University of Sport, Taichung City, Taiwan;
Department of Management Information Systems, National Chung Hsing University, Taichung City, Taiwan

Abstract: The use of Electroencephalography (EEG) signal combining perceived stress scale to conduct long-term stress analysis has recently become an important field of re-search. Yet, how to analyze the physiological signals from the brain is an im-portant issue for increasing the accuracy of stress assessment. The purpose of this study is to identify an optimal stress classification model by analyzing, with densely connected convolutional network methods, the data on players' EEG that are collected using perceived stress scale. Research method involves four stages: data collection, pre-processing, feature extraction, and classification. Our findings suggested that two-level classification for stresses had an accuracy at 98.35%. We provided the optimal classification results, which makes this study distinct from past studies.

Statistical Criteria for Gene Selections

Hong Zhang
Georgia Southern University, Savannah, Georgia, USA

Abstract: Gene selection is an important processing step in machine learning based analysis of gene expression data produced by genomic testing technologies such as microarray and next-generation sequencing. We propose two statistical criteria to measure the relevance of genes for the classification problem. The measures address the problems of accuracy and robustness of the selection process and will lead to an efficient, effective, and robust gene selection and ranking algorithm.

Detecting the Most Relevant Slices in MRI Parkinson Disease for Multi-Classification using Evolutionary Algorithm

Olga Valenzuela, Francisco Carrillo-Perez, Luis Javier Herrera, Hector Pomares, Fernando Rojas, Ignacio Rojas
Faculty of Science, University of Granada, Spain;
Department of Computer Architecture and Technology, University of Granada. C.I.T.I.C.,
Periodista Rafael Gomez Montero, Granada, Spain

Abstract: The correct detection of the Parkison's disease is crucial for providing the patient with an effective treatment and for improving the prognosis. Creating clinical decision support systems that efficiently classify this disease from other forms of parkinsonims would be of great help for clinicians. We present a methodology for selecting the best slices in a MRI for building accurate classification of five different classes (Normal, SWEDD, Prodromal, PD and Gen Cohort). An evolutionary algorithm-based system is used to select the best slices, and an SVM is used as the classification model. The model obtains an accuracy of 90\% with five slices in the testing phase.

The 8th International Conference on Biomedical Engineering & Sciences (BIOENG'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-BIOENG
https://american-cse.org/csce2022/

Challenges Facing Traditional Medicine

Matthew N. O. Sadiku, Janet O. Sadiku, Olaniyi D. Olaleye, Sarhan M. Musa
Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, Texas, USA;
Barbara Jordan-Mickey Leland School of Public Affairs, Texas Southern University, Houston, Texas, USA;
Department of Manufacturing Engineering, Central State University, Wilberforce, Ohio, USA

Abstract: Traditional medication essentially entails the use of herbal remedies, animal parts, and minerals. It includes a diversity of health practices, approaches, knowledge, spiritual therapies, manual techniques, and exercises, applied to maintain well-being through treating, diagnosing, or preventing illnesses. The increasing use and popularity of traditional medicines have created challenges in public health from the point of view of politics, safety, efficacy, quality control, access, and rational use. A common criticism of traditional medicine is that medical doctors treat symptoms instead of looking for the root cause of the symptoms. The current challenge is to pursue action along three lines: evaluation, integration, and training. This paper addresses twenty challenges facing the practice of traditional medicine all over the world.

Comparison of Deep Learning Architectures for COVID-19 Multi-class Classification over CT-Scans

Jesus Toledano Pavon, Francisco Carrillo-Perez, Juan Carlos Morales Vega, Luis Javier Herrera, Ignacio Rojas
Department of Computer Architecture and Technology, University of Granada. C.I.T.I.C.,
Periodista Rafael Gomez Montero, Granada, Spain;
Stanford Center for Biomedical Informatics Research (BMIR), Department of Medicine,
Stanford University, Stanford, California, USA

Abstract: COVID-19 disease continues to spread with the appearance of new variants of the virus in the current year 2022. For this reason different deep learning methods have been tested for COVID-19 prediction. In this paper, a multi-class classification between COVID-19, Community Acquired Pneumonia (CAP) and normal patients was carried out for multiple Convolutional Neural Networks (CNN) models by using labelled scans of Computed Tomography (CT). Our main contribution involves a custom pre-processing algorithm which removes noise and artifacts around patient\'s body enhancing the input slices of CNN models. After applying a 5-fold cross validation to the training phase for each model, VGG16 turned out to be the best model with 90% ± 1.22 test accuracy at slice-level and 99% ± 1.19 test accuracy at patient-level. Thus, the proposed method is able to distinguish COVID-19 patients from CAP and normal patients quite precisely.

CNN Performance of Node Numbers at FC Layer in sEMG Signals based Finger Number Recognition

Sung-Hwan Bae, Chun-Ki Kwon
Soonchunhyang University, Asan, ChungNam, South Korea

Abstract: This work investigates appropriate numbers of layer and its nodes at fully connected layers for CNN based finger number recognition using sEMG signals. To the end, number of nodes at each FC layer is selected by reducing the number of nodes by half times as many as number of nodes at the previous layer, starting at 512 nodes. Appropriate number of layers and nodes at FC layers in CNN is determined to show best recognition rate of five Korean finger number gestures from (1) to five (5). The results of the performed experiment show that three FC layers with 512, 256, and 128 nodes achieved the best recognition rate of 96% in CNN-based Korea finger number recognition.

Deep Learning-Based MR Image Re-parameterization Abhijeet Narang, Abhigyan Raj, Mihaela Pop, Mehran Ebrahimi

Indian Institute of Technology (ISM), Dhanbad, India;
Sunnybrook Research Institute, Toronto, Ontario, Canada;
Faculty of Science, Ontario Tech University, Oshawa, Ontario, Canada

Abstract: Magnetic resonance (MR) image re-parameterization refers to the process of generating an MR image with a new set of MRI scanning parameters using simulations. Different parameter values generate distinct contrast between different tissues, helping identify pathologic tissue. Typically, more than one scan is required for diagnosis; however, acquiring repeated scans can be costly, time-consuming, and difficult for patients. Thus, using MR image re-parameterization to predict and estimate the contrast in these imaging scans can be an effective alternative. In this work, we propose a novel deep learning (DL) based convolutional model for MRI re-parameterization. Based on our preliminary results, DL-based techniques hold the potential to learn the non-linearities that govern the re-parameterization.

An Experimental Comparison of Voxel Processing Algorithms for Automatic Cephalometric Landmarking on CBCT Volumes

Jesus Montufar Trujillo, Ernesto Hernandez Rosales, Rogelio J. Scougall-Vilchis
Universidad Politecnica de Otzolotepec, Mexico;
Universidad Autonoma del Estado de Mexico, Mexico

Abstract: Cephalometry is a diagnostic method in orthodontics to ob- tain cranial measurements, based on location of cephalometric land- marks, recently can be performed on CBCT (Cone Beam Computed Tomography) volumes. This paper presents an algorithm for automatic localization of cephalometric landmarks on CBCT volumes and an ex- perimental comparison versus two state of the art techniques. We report the automatic location of 18 common cephalometric landmarks on 24 CBCT volumes obtained from a public database. Localization results were evaluated by linear measurements compared to the manually an- notated groundtruth. Localization time was estimated based on each al- gorithm complexity and its CPU processing time. Our new algorithm scored fast automatic landmark localization, near to a real-time perfor- mance; on the other hand, the accuracy in localization reports a 2.4 mm mean error. We conclude that our method performs a quick and enough accurate initial search based on 2D volume projections and saves compu- tational time despite processing complete volumes for 3D cephalometry.

The 20th International Conference on Scientific Computing (CSC'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-CSC
https://american-cse.org/csce2022/

Educational Data Mining Using Python and Machine Learning Algorithms

May Alsofyani, Ali Jaber Almalki, Liqiang Wang
Department of Computer Science, University of Central Florida, Orlando, Florida, USA

Abstract: Extracting relevant information from educational data is difficult as there is a huge amount of redundant data. Irrelevant attributes in the data make it difficult to gather the accuracy of the data. Educational Data mining helps to extract the hidden features of the data. Using the machine learning algorithm in Educational Data Mining extracts the accuracy of the data without changing the original attributes of the data. In this research paper Educational Data Mining is performed using PYTHON and WEKA. The data in which the algorithms are performed are taken from Kaggle.com. The comparison results are performed by comparing the results from Python and Weka. The importance of this research is to improve the accuracy of Educational Data.

Maximum Match Subsequence Alignment Algorithm - Finely Grained (MMSAA - FG)

Bharath Reddy, Richard Fields
Process Automation R&D, Schneider-Electric, Lake Forest, California, USA

Abstract: Sequence alignment is common nowadays as it is used in many fields to determine how closely two sequences are related and at times to see how little they differ. In computational biology / Bioinformatics, there are many algorithms developed over the course of time to not only align two sequences quickly but also get good laboratory results from these alignments. The first algorithms developed were based of a technique called Dynamic Programming, which were very slow but were optimal when it comes to sensitivity. To improve speed, more algorithms today are based of heuristic approach, by sacrificing sensitivity. In this paper, we are going to improve on a heuristic algorithm called MASAA (Multiple Anchor Staged Local Sequence Alignment Algorithm) . and MASAA -S which we published previously. This new algorithm appropriately called MMSAA . FG, stands for Maximum Match Subsequence Alignment Algorithm . Finely Grained. The algorithm is based on suffix tree data structure like our previous algorithms, but to improve sensitivity, we employ adaptive seeds, and finely grained perfect match seeds in between the already identified anchors. We tested this algorithm on a randomly generated sequences, and Rosetta dataset where the sequence length ranged up to 500 thousand.

Analysis of all Local Pairwise Sequence Alignment Algorithms

Bharath Reddy, Richard Fields
Process Automation R&D, Schneider-Electric, Lake Forest, California, USA

Abstract: Biological sequence alignment is common today and are used in a variety of fields ranging from Bioinformatics, Computational Biology, Genome analysis, Cancer research, Stem Research and many more fields. Most of these fields use the sequence alignment to find the .similar. regions or similarities between organisms. Since, this step is computational heavy, today, there are specialized hardware to help speed up and techniques and strategies to help speed up or improve the sensitivity (quality) of the alignment in general. The early successful algorithms in sequence alignment were focused on quality, and it produced an optimal algorithm called SmithWaterman algorithm, which we will discuss in detail later using a technique called .Dynamic Programming.. The time complexity of this algorithms was O (mn). Later, to speedup, heuristic algorithms were developed. Heuristic algorithms gave up a little bit on the quality for speed, by calculating the near-optimal alignment rather than optimal algorithm. In this paper, we will analyze various computational approaches for local sequence alignments.

Towards AI-based Audio Processing: A Literature Review

Vladislav Rykov, Praveen Meduri
California State University, Sacramento, California, USA

Abstract: This paper describes the shift from conventional audio signal processing methods to AI-based methods, with focus on generative models, specifically Generative Adversarial Networks (GANs). Conventional methods are described first, building a theoretical background for modern techniques. Then, Neural Networks (NN) based methods are discussed, followed by a discussion on deep models, Variational Autoencoders (VAE) and GANs.

Photogrammetry Software

Daniel L. Bowen, Tina V. Johnson, Catherine V. Stringfellow
Department of Computer Science, Midwestern State University Texas, Wichita Falls, Texas, USA

Abstract: As virtual reality gains momentum, the need to efficiently render 3D environments has become increasingly important. Historically, getting a 3D scan of a room has required LiDAR equipment as well as a 360-panorama camera with complex interpolation and stitching to create a rather rough image. Advances have been made, both academically and commercially, in the ability to use 2D images to create 3D models of a location. This paper provides an overview of two open source photogrammetry software packages, COLMAP and AliceVision Meshroom.

Data-aware Customization of Activation Functions Reduces Neural Network Error

Fuchang Gao, Boyu Zhang
Department of Mathematics and Statistical Science, University of Idaho, Idaho, USA;
Institute for Modeling Collaboration and Innovation, University of Idaho, Idaho, USA

Abstract: Activation functions play critical roles in neural networks, yet current off-the-shelf neural networks pay little attention to the specific choice of activation functions used. Here we show that data-aware customization of activation functions can result in striking reductions in neural network error. We first give a simple linear algebraic explanation of the role of activation functions in neural networks; then, through connection with the Diaconis-Shahshahani Approximation Theorem, we propose a set of criteria for good activation functions. As a case study, we consider regression tasks with a partially exchangeable target function, i.e. f(u,v,w)=f(v,u,w) for u, in {R}^d and w in {R}^k, and prove that for such a target function, using an even activation function in at least one of the layers guarantees that the prediction preserves partial exchangeability for best performance. Since even activation functions are seldom used in practice, we designed the ``seagull'' even activation function log(1+x^2) according to our criteria. Empirical testing on over two dozen 9-25 dimensional examples with different local smoothness, curvature, and degree of exchangeability revealed that a simple substitution with the ``seagull'' activation function in an already-refined neural network can lead to an order-of-magnitude reduction in error. This improvement was most pronounced when the activation function substitution was applied to the layer in which the exchangeable variables are connected for the first time. While the improvement is greatest for low-dimensional data, experiments on the CIFAR10 image classification dataset showed that use of ``seagull'' can reduce error even for high-dimensional cases. These results collectively highlight the potential of customizing activation functions as a general approach to improve neural network performance.

Design of Multi-exaflops Machines using Equality Network Topology

Chi-Hsiu Liang, Chun-Ho Cheng, Yng-Ching Wu, Chao-Chin Li, Chun-Ming Chen, Po-Lin Huang, Chi-Chuan Hwang
Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan

Abstract: A high-performance chordal-ring interconnect topology system named Equality is revisited in this paper. Equality interconnects are highly symmetric. This paper details the procedures for the construction of an Equality interconnect, and the evaluation of its performance using open-source BookSim package. A 16,384-endpoint system is formulated in this study to compare with a few popular network topologies to show the capability of Equality under the same hardware constraints. The results show that the Equality networks are resilient under various common traffic models. This paper also demonstrates that a million-endpoint system is feasible with Equality network with a single-point cycle-accurate simulation.

Transmission Range Test for a LoRa-based in-situ Water Assessment System in the Vaal Region

Christian N. Ejike, Tebello Mathaba, George Ochieng, Marcel O. Odhiambo
Department of Electrical Engineering, Vaal University of Technology, Vanderbijlpark, South Africa;
Postgraduate School of Engineering Management, University of Johannesburg, South Africa;
Faculty of Engineering, Mangosuthu University of Technology, Durban, South Africa

Abstract: Recent WSN and IoT studies have illustrated the application of LoRa and LoRaWAN as a viable Low-Power Wide Area Network (LPWAN) technology for in-situ monitoring scenarios. An evaluation of the use of LoRa and LoRaWAN as a means of communication for in-situ water assessment over a distance of about 2 km was conducted and the results are presented. The experimental tests focus on the use of different LoRa spread factor (SF) between 7 to 12 across four locations to examine its impact on the packet delivery ratio (PDR) within the same transmission ranges. The results are discussed, and the objectives set in this study were met.

A Machine Learning Approach for Targeting Network Anomalies with an Intrusion Detection System

Haydar Teymourlouei, Lethia Jackson, Daryl Stone
Department of Technology & Security, Bowie State University, Bowie, Maryland, USA

Abstract: Network anomaly detection systems allow monitoring of computer networks that perform differently from the network protocol. They can be implemented in various domains. Working on the internet is a critical task, due to the many network security threats, like intrusions. It is essential to recognize the challenges and identify solutions to reduce the attacks. In this paper, Intrusion Detection Systems (IDS) will be researched and evaluated as a solution to deal with attacks in various environments. We conduct an experiment using supervised Machine Learning (ML) for a network anomaly detection system that will provide real-time anomaly detection in network time-series data, data collection, and training modules. Machine learning algorithms are used to analyze unusual instances in a particular network. These methods proved to be an efficient way to detect network anomalies while meeting the specified system performance.

An Intelligent Smart Parking System Leveraging on IoT Parking Sensors and AI-based Prediction Algorithms for Parking Occupancy Optimization

Elena Politi, George Dimitrakopoulos, Eleni Giannopoulou, Michalis Mitrou, Alexandros Diakogiannis, Vera Stavroulaki, Panagiotis Demestichas, Alexandros Zacharis
Department of Informatics and Telematics, Harokopio University of Athens, Greece;
WINGS ICT Solutions, Athens, Greece; NETLINK, Athens, Greece

Abstract: The Internet of Things (IoT) can be considered as the next big technological thing, where novel technologies are being introduced to modernize different aspects of modern life and especially the smart city ecosystem. In particular, IoT solutions can address the problems of citizens in dense urban areas, such as the problem finding an available parking spot. Intelligent management parking systems offer the benefits of sparing a significant amount of time to the driver, while improving traffic congestion and reducing fuel consumption of the vehicle. This work proposes a framework for an intelligent parking system based on IoT technology and AI-based prediction algorithms for parking occupancy optimization. The efficiency of our proposed solution is showcased through three use case scenarios based on real data obtained by parking sensors.

Mitigating the Effects of Attribute Noise Using Two-Level Filtering Learner Algorithm

Eman Allogmani, Darsana Josyula
Department of Computer Science, Bowie State University, Bowie, Maryland, USA;
Department of Computer Science, Majmaah University, Al-Majmaah, Saudi Arabia

Abstract: Noise imposes negative impacts on learning algorithms by misleading the training and thus causing erroneous predictive classifications. Attribute noise occurs in the values of the independent variables used to predict dependent variables. This paper discusses an efficient technique for mitigating the impact of noise and relies on identifying and filtering outliers during the training process in order to enhance the resulting classification model. This technique, which we refer to as the two-level filtering learner algorithm, is applied to two types of attribute noise: domain and Gaussian attribute noise, and the results of these experiments, which we present in this paper, show the feasibility of our method in creating predictive models with improved classification accuracy.

ASA: A Simulation Environment for Evaluating Military Operational Scenarios

Joao P. A. Dantas, Andre N. Costa, Vitor C. F. Gomes, Andre R. Kuroswiski, Felipe L. L. Medeiros, Diego Geraldo
Decision Support Systems Subdivision, Institute for Advanced Studies, Sao Jose dos Campos, Sao Paulo, Brazil

Abstract: The Aerospace Simulation Environment (Ambiente de Simulação Aeroespacial -- ASA in Portuguese) is a custom-made object-oriented simulation framework developed mainly in C++ that enables the modeling and simulation of military operational scenarios to support the development of tactics and procedures in the aerospace context for the Brazilian Air Force. This work describes the ASA framework, bringing its distributed architecture for managing multiple simulation machines, a data analysis platform for post-processing simulation data, the capability of loading models at simulation runtime, and a batch mode execution platform to perform multiple independent executions simultaneously. In addition, we present a list of recent works using the ASA framework as a simulation tool in the air combat context.

Driver Safety Reward with Cooperative Platooning using Blockchain

Sruthi Rachamalla, Henry Hexmoor
School of Computing, Southern Illinois University, Carbondale, Illinois, USA

Abstract: Cooperative driving (or Platooning) focuses on improving the safety and efficiency by connecting two or more vehicles on a road by vehicular communication protocols. The leader is crucial as it manages the platoon, establishes communication between cars, and performs platoon maneuvers. In this paper, we proposed a driver incentive model which encourages platooning on roads leading to driver safety. As, the leader of platoon have multiple responsibilities than followers, our model rewards more incentives to leader than followers. These incentives will be rewarded as crypto tokens. This digital monetization method for both leaders and followers of a platoon is accomplished by secure transactions using blockchain.

A Tele-Monitoring of Chronic Heart Failure Patients using Wireless Body Area Networks: Data Transmission Time, Throughput and Reliability

Ajay Thapa, Abeer Alsadoon, Muhammad Siddiqi, Benoy Varghese, Fatima Furqan,
MD Rafiqul Islam, Shahad Ahmed, Shayma Ismail Ali
Charles Sturt University, Australia; Western Sydney University, Australia; Kent Institute Australia, Australia;
Asia Pacific International College, Australia; The University of Duhok, KRG, Iraq

Abstract: In tele-monitoring, it is challenging to transfer chronic heart pa-tients. data in real-time successfully due to the limitations of emergency data loss by maximum data overflow and interference in WBANs. This paper fo-cuses on improving transmission time, throughput and reliability of data dur-ing data transmission in WBANs. The proposed system comprises Priority-based Data Transfer and Interference-aware Traffic-priority-based Link Scheduling (PDTaITLS) to reduce the loss of data and interference in data transmission. This improves transmission time, network throughput and reli-ability of data transmission by listing sensor nodes into interfered and non-interfered groups. Afterward, each group.s emergency data is transferred at the same time with first-time slots as per priority basis. The results show that the new enhanced technique achieves improvements in emergency data transmission time, network throughput and data reliability.

Antecedents of Performative Action from Social Media Exchanges

Michael D. Workman, Michael Councilman
Thomas School of Business, University of North Carolina at Pembroke, North Carolina, USA

Abstract: Opensource intelligence (OSINT) is used to gather information from social media and other open sources, while deep fakes, fake news and other forms of extreme expression in social media are designed to misdirect or to disseminate misinformation or disinformation. In our previous research, we found that some people were more susceptible than others to influences from social media, along with degrees of influence as shown in the affect intensity associated with their sentiments. In our follow-on study, we used deep machine learning (unsupervised and supervised) to (1) identify characteristics associated with those who may be influenced (positively or negatively) by social media, and (2) from that, determine who might be most likely to take increasingly aggressive forms of action. In this present study, we used structured equation modeling (SEM) of concomitant participant responses to a self-report questionnaire to triangulate our previous machine learning findings, to develop a theoretical threat model of the main factors for generalization and further empirical testing in other contexts.

How to Extend the Abstraction Refinement Model for Systems with Emergent Behavior?

Mohamed Toufik Ailane, Christoph Knieke, Andreas Rausch
Clausthal University of Technology, Institute for Software & Systems Engineering, Clausthal-Zellerfeld, Germany

Abstract: The Abstraction Refinement Model has been widely adopted since it was firstly proposed many decades ago. This powerful model of software evolution process brings important properties into the system under development, properties such as the guarantee that no extra behavior (specifically harmful behaviors) will be observed once the system is deployed. However, perfect systems with such a guarantee are not a common thing to find in real world cases, anomalies and unspecified behaviors will always find a way to manifest in our systems, behaviors that are addressed in this paper with the name ``emergent behavior\''. In this paper, we extend the Abstract Refinement Model to include the concept of the emergent behavior. Eventually, this should enable system developers to: (i) Concretely define what an emergent behavior is, (ii) help reason about the potential sources of the emergent behavior along the development process, which in return will help in controlling the emergent behavior at early steps of the development process.

Numerical Diffusion of Material in Numerical Simulations for Hydrodynamics

William W. Dai, Jim M. Ferguson
Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Abstract: This paper is to investigate the numerical diffusion of material in numerical simulations through a computer code, xRAGE [1], in which a few different numerical approaches for hydrodynamics are used, particularly, typical dimensionally split and unsplit algorithms, techniques of interface steepening for material interface, and material interface reconstruction. Without any treatment for material interface an interface of material could be numerically diffused over around 35 cells. Even with adaptive mesh refinement, numerical simulations through Eulerian codes are very diffusive for material interface. With the treatment of interface steepening, the spreading could be cut by a factor 2. Although the techniques of interface steepening significantly reduce numerical diffusion of material, the simulations with these techniques are still rather diffusive. In practice it is hard to minimize numerical diffusions of material through refinement of mesh without reconstruction of material interface. The reconstruction of material interface is the only approach that could effectively minimize the numerical diffusion of material. If applicable, the reconstruction of material interface is strongly recommended together with the refinement of mesh for problems in which numerical diffusion of material could be a significant source of error in numerical simulations.

WalkOrder: Combining Graph Sampling and Reordering for Fast Computation

Kohei Tsuchida, Kunitake Kaneko
Graduate School of Science and Technology, Keio University, Yokohama, Japan;
Faculty of Science & Technology, Keio University, Research Institute for Digital Media Content, Keio University, Japan

Abstract: A smaller graph sampled from a large graph uses an excessively wide range of vertex IDs beyond its graph size, which causes a sparse layout on memory. Gorder is a state-of-the-art graph reordering technique for layout optimization. However, despite providing performance improvement, applying Gorder to a sampled graph requires lengthy and impractical time periods, such as reordering time and waiting time for the completion of sampling. In addition, it is unsuitable for multicore graph computation because high and low degree vertices are separated on memory by reordering. We propose WalkOrder, a technique that conducts gradual graph reordering in parallel with sampling; sampling is performed by the widely used Metropolis-Hastings Random Walk. WalkOrder reduces reordering costs of Gorder and balances workload between cores. Our extensive experiments on six real-world graphs revealed that WalkOrder was faster than baseline techniques by an average of 1.52 times and a maximum of 1.95 times.

Behavioral Player Rating in Competitive Online Shooter Games

Arman Dehpanah, Muheeb Faizan Ghori, Jonathan Gemmell, Bamshad Mobasher
School of Computing, DePaul University, Chicago, Illinois, USA

Abstract: Competitive online games use rating systems for matchmaking; progression-based algorithms that estimate the skill level of players with interpretable ratings in terms of the outcome of the games they played. However, the overall experience of players is shaped by factors beyond the sole outcome of their games. In this paper, we engineer several features from in-game statistics to model players and create ratings that accurately represent their behavior and true performance level. We then compare the estimating power of our behavioral ratings against ratings created with three mainstream rating systems by predicting rank of players in four popular game modes from the competitive shooter genre. Our results show that the behavioral ratings present more accurate performance estimations while maintaining the interpretability of the created representations. Considering different aspects of the playing behavior of players and using behavioral ratings for matchmaking can lead to match-ups that are more aligned with players' goals and interests, consequently resulting in a more enjoyable gaming experience.

ARIMA Model Analysis for Portfolio Selection

Pei-Chun Lin, Nureize Arbaiy, Shamshul Bahar Yaakob, Qu-Kai Lin, Ching-Hung Chiu
Department of Information Engineering & Computer Science, Feng Chia University, Taichung, Taiwan;
Faculty of Computer Science & Information Technology, Universiti Tun Hussein Onn Malaysia, Malaysia;
Fakulti Teknologi Kejuruteraan Elektrik, Universiti Malaysia Perlis, Malaysia

Abstract: In early 2020, crown pneumonia (COVID-19) began to spread worldwide. It has a huge impact on the daily lives of people around the world. The outbreak has had a major impact on our business operations and share prices this year. Many countries and cities have implemented long-distance work laws as a result of the outbreak to continue their business operations. Companies that operate their factories sometimes are at risk of product shortages due to declining production and limited factory operations. The weakening global semiconductor supply scenario also increased TSMC stock. However, not all stocks gained in value, and many investors lost a lot of money. Accordingly, forecasting future trends is critical in assisting decision-makers. This paper is aimed to utilize time series analysis (Auto-Regressive Integrated Moving Average) to forecast the stock market and then develop a preferred model that can let investors make the decisions in time and choose the best investment strategy. We use the yahoo stock market database in this paper. The data contains daily transaction volume, opening price, highest price, lowest price, closing price, price difference, and numbers of transactions for the previous ten years. The portfolio selection was made up of shares from five different companies. This strategy relies on a good ARIMA model by data pre-processing and modeling. Investors can predict the overall direction of the stock market for next year and reduce losses caused by economic effects by using our ARIMA model.

Differentially Private Synthetic Health Data

Joseph Pedersen, Kristin P. Bennett
Rensselaer Polytechnic Institute, Troy, New York, USA

Abstract: Generative adversarial networks (GANs) have been successfully used to generate synthetic health data to preserve patient privacy. We develop a differential private (DP) Wasserstein GAN with gradient penalty (WGAN-GP) that empirically demonstrates high utility, resemblance, and privacy on three health care problems. We derive tight bounds on the gradient of the loss for a WGAN for certain architectures. Computational tests show the noise added to achieve DP results in convergence to better local minima. The proposed DP WGAN-GP generates synthetic data that performs better in classification tasks than the original WGAN with no noise. We hypothesize that adding differential privacy reduces the problem of mode collapses that frequently occurs in GANs, and perhaps the regularization that it induces may benefit other optimization problems.

Ae^2I: A Double Autoencoder for Imputation of Missing Values

Fuchang Gao
Department of Mathematics and Statistical Science, University of Idaho, Moscow, Idaho, USA

Abstract: The most common strategy of imputing missing values in a table is to study either the column-column relationship or the row-row relationship of the data table, then use the relationship to impute the missing values based on the non-missing values from other columns of the same row, or from the other rows of the same column. This paper introduces a double autoencoder for imputation (Ae^2I) that simultaneously and collaboratively uses both row-row relationship and column-column relationship to impute the missing values. Empirical tests on Movielens 1M dataset demonstrated that Ae^2I outperforms the current state-of-the-art models for recommender systems by a significant margin.

Network Connectivity Knowledge Graph using DWave Quantum Hybrid Solvers

Nivedha Rajaram
UST-GLOBAL, USA

Abstract: Hybrid Quantum solvers are given prime focus in recent years by computation problem solving domain industrial applications. D'Wave Quantum Computers are one such paragon of systems build us- ing quantum annealing mechanism. Discrete Quadratic Models is a hy- brid quantum computing model class supplied by D'Wave Ocean SDK - a real time software platform for hybrid quantum solvers . These hybrid quantum computing modellers can be employed to solve classic problems. One such problem that we consider in this paper is to find network con- nectivity knowledge hub in huge network of systems. Using this quantum solvers we try to find out the prime system hub which act as a supreme connection points for the set of connected computers in a large network. This paper establishes an innovative problem approach to generate con- nectivity system hub plot for a set of systems using DWave ocean sdk hybrid quantum solvers.

Tridiagonal Matrix Inversion on the GPU

Peter Yoon
Department of Computer Science, Trinity College, Hartford, Connecticut, USA

Abstract: Inverting a matrix is more computationally challenging than solving a linear system. However, in fields such as structural engineering, dynamic systems, and cryptography, computing the inverse of a matrix is often necessary. This paper presents an accelerated procedure for computing the inverse of diagonally dominant tridiagonal matrices on the GPU. The algorithm is based on the recursive application of the Sherman-Morrison formula for tridiagonal matrices. The preliminary results show that our GPU implementation of the inversion procedure outperforms the conventional CPU-based implementations with a speedup of up to 24x.

Understanding Health Information Overload in Patients Using Technology in the UAE Amidst COVID-19

Naja Yoosuf, Zeenath Khan
Department of Engineering & Information Science, University of Wollongong in Dubai, Dubai, United Arab Emirates

Abstract: The amount of health information is growing day by day to an extent that medical professionals have to read 5000 articles per day to keep track of the latest medical research. Information overload is when the information acquired by an individual within a time duration exceeds their ability to comprehend the information. Previous studies had not generalized the overall effects of overload and thus this research aims to do so by analyzing the impact of information overload in UAE residents during COVID-19. This research portrays that health information overload is present in the UAE and that fatigue occurs because of it. It also indicates that anxiety could alter the fatigue extent felt due to overload while motivation cannot.

The New Clustering Organization Enabling Fast Search of High Dimensional Big Data

Zohreh Safari, Yu Zhuang
Department of Computer Science, Texas Tech University, Lubbock, Texas, USA

Abstract: This paper presents a clustering data set organization for high-dimensional search spaces. The method is designed for similarity searches where one would like to find many data points close to a target query more efficiently. It has great attention for many search and detection applications like similar image/text search and document detection. Our scheme focuses on indexing high dimensional data, creating clusters without the limitation of the pre-defining number of clusters, and then searching near data points in a simple but efficient approach. Our organization data set is called Bisecting Radius K-Means algorithm (BR-K-Means). BR-K-Means enables the efficient search of high dimensional big data that compares favorably in experiments with probably the best existing tree-based structure in total computation time, including data organization and data searching. The results indicated that our linear radius-based clustering approach could search closer data points faster than tree structures in high-dimensional big data.

Evaluation of Reduced-precision Computation on the Tendral Statistical Interpolation System (T-SIS) for Ocean Modeling

Daniel Voss, Gary Tyson, Ashwanth Srinivasan
Computer Science, Florida State University, Tallahassee, Florida, USA;
Tendral, LLC, Key Biscayne, Florida, USA

Abstract: Geophysical circulation models use double- and single-precision to represent numbers for high precision arithmetic. Recent research efforts have focused on using reduced or mixed-precision in a sample of these models to speed up the simulations while maintaining numerical accuracy. Data assimilation techniques are an ideal field for testing the effect of reduced-precision because they are inherently prone to observational error, variable resolution, and interpolation of data on model grids. The amount and complexity of observational data is increasing rapidly and it is important to explore techniques to reduce the time and increase the computational performance of data assimilation methods. This paper presents the results of evaluating reduced-precision on the unique implementation of the Tendral Statistical Interpolation System (T-SIS), a Gaussian Markov Random Fields (GMRF) data assimilation scheme for ocean modeling.

Trends and Strategies to Optimize Training Processes to Decrease Learning Time in Deep Neural Networks

Miroslava Barua, Patricia A. Nava
Department of Electrical & Computer Engineering, The University of Texas at El Paso, El Paso, Texas, USA

Abstract: Artificial intelligence implements deep learning solutions using big data coupled with large networks that require long and costly training times. Deep Learning (DL) networks are composed of many units interconnected in many layers creating structures with multiple parameters and numerous connections that are constantly updated through the standard learning process. Training a DL network is a long process which includes evaluating the DL network's final response, and iterative-ly changing parameters, such that the selected cost function is minimized. In this paper we take a look at the trends and strategies that have been used to optimize the training of deep neural networks and how the incorporation of fuzzy logic can shorten the learning process. In our on-going research, we will explore these ideas and will investigate the use of a fuzzy decision support system that could determine when is beneficial to invoke the backpropagation learning rules and complete the process of updating the neuronal weights during the learning process used in deep neural networks.

WORKSHOP/SESSION
Military and Defense Modeling and Simulation

Co-Chairs: Prof. Douglas D. Hodson (Chair), Prof. Michael R. Grimaila**, Dr. Richard Dill**
**Computer Science and Engineering Department, US Air Force Institute of Technology (AFIT), USA*
***Head, Systems Engineering & Management Department, US Air Force Institute of Technology (AFIT), USA*

Integrating Data Distribution Service (DDS)-Cerberus into Cyclone

Jesse Sidhu, Richard Dill, Wayne Henry, Douglas Hodson
US Air Force Institute of Technology (AFIT), USA

Abstract: Eclipse's Cyclone Data Distribution System (Cyclone DDS) lacks built-in authentication processes. This lack of built-in authentication makes Cyclone DDS vulnerable to rogue nodes. Data Distribution System - Cerberus (DDS-C) securely authenticates publishers and subscribers using Kerberos ticketing. We implement DDS-C into Cyclone DDS making a plugin for Cyclone DDS called Cyclone DDS-C. Our implementation of Cyclone DDS-C utilizes both Kerberos and Cyclone DDS libraries. An evaluation measures both system latency and throughput to derive differences between baseline Cyclone DDS and two differing Cyclone DDS-C implementations. One implementation uses Upfront Authorization and the other uses Continuous Authorization of Kerberos authentication methods. Experimental results show that Cyclone DDS-C using Upfront Authorization has a 17.12% average increase in latency and 4.32% average decrease in throughput compared to Cyclone DDS. Cyclone DDS-C using Continuous Authorization has an 80.13% average increase in latency and 79.72% average decrease in throughput abilities as compared to Cyclone DDS. This research provides a means to add authentication security in Cyclone DDS.

Benchmarking Performance of Unity's Data Oriented Technology Stack

Brett M. Martin, Michael J. Visci, John W. Visci, Jonathan Thompson, Douglas Hodson
US Air Force Institute of Technology (AFIT), USA

Abstract: N/R

SPARK and Rust: An Overview of Safe and Reliable Software

Brian S. Wheelhouse, Kenneth M. Hopkinson, Laura R. Humphrey, Douglas Hodson
US Air Force Institute of Technology (AFIT), USA

Abstract: It is never a simple task to create highly reliable and secure software. Software testing only proves that software errors exist, not that software is error-free. Additionally, most programming languages seem to have only taken into account reliability and security features as an afterthought, leaving the burden on the software developer to have the knowledge and care to ensure the security and reliability of the software they produce. As software has become increasingly complex, some programming languages have begun to consider reliability and security as key elements of the language design itself. Two languages, SPARK (a subset of Ada) and Rust are prime examples of this trend in language design. Both languages also have software verification tools available that allow developers to prove the software is functionally correct according to its specifications and that it is free from runtime errors and other possible bugs through static analysis. This paper provides an overview of these two languages and what makes them safe and reliable with a particular emphasis on the SPARK language and verification toolset.

The 21st International Conference on e-Learning, e-Business, Enterprise Information Systems, & e-Government (EEE'22: July 25-28, 2022, USA)

https://american-cse.org/csce2022/conferences-EEE

https://american-cse.org/csce2022/

Determinants of Behavioral Intention to e-Learning during the COVID-19: System Factors and Risk Perception

Hyeon Jo
Department of Strategic Planning, RealSecu, Busan, South Korea

Abstract: Since the outbreak of COVID-19, the use of e-learning has significantly increased. Now, e-learning performs not only an educational role but also a health function that reduces social contact. This study identifies the factors that impact behavioral intention to use e-learning under COVID-19. For empirical analysis, a total of 360 samples were collected by surveying university students in South Korea and Vietnam. This research validated the model using the partial least squares–structural equation modeling (PLS-SEM) technique. The result of the study showed that behavioral intention is affected by system interactivity, system response, and affective risk perception. The findings of the current study would help sustainable education in the COVID-19 crisis.

Healthcare Information Security: A Paradigm Shift to Zero Trust Principles

David Ang, Onome Christopher Edo
Department of Information Systems, College of Business, Auburn University at Montgomery, Alabama, USA

Abstract: Information security threat continues to ravage enterprises over the world, and the healthcare industry is not left out of this chain, patient healthcare record is sensitive, and a foolproof security measure is required to safeguard these assets. Security frameworks have been advanced to counter this threat, but hackers continue to find loopholes within and outside the information system perimeters. The purpose of this study was to advance a security framework based on Zero Trust Architecture (ZTA). The hypothetical framework incorporates a policy engine, policy administrator and policy enforcement point which are the base components of the zero-trust architecture, founded on the principle of .trust no one. the model seeks to authenticate and validate users and devices continuously. Through the proposed framework, patient data can be retrieved and updated, and the threat of intrusion can be curtailed while fostering interoperability of patient data.

The Evaluation of Mobile Technology Adoption as an Employee Training Tool during the Pandemic

Anastasia Tracy Biggs
School of Computer Science and IT, Colorado Technical University, Colorado Springs, Colorado, USA

Abstract: The purpose of this qualitative case study was to explore why and how corporate training managers can adopt mobile technology for employee training. The case study method explored the learning processes to determine if a learning model is appropriate for the use of mobile technology as a training tool (De Zan, De Toni, Fornasier, and Battistella, 2015, p. 341). This qualitative case study utilized interviews to explore how the use of mobile technology can be adopted to train employees. The use of interviews examined the degree of employee growth from mobile training (Alberghini, Cricelli, and Grimaldi, 2014, p. 260). Case study methodology answered how mobile technology through cause-effect relationships explored the lack of mobile technology adoption interventions between corporate managers and the organization (De Zan et al, 2015, p. 335) (Tsang, 2013, p. 197).

Automatic Dynamic and Adaptive Playout of Competency-Based Learning Units Based on Data in Learners' Competency and Qualification Profile

Ramona Srbecky, Michael Winterhagen, Andre Hedderoth, Ivo Ochsendorf, Felix Fischman, Matthias Then, Benjamin Wallenborn, Binh Vu, Wieland Fraas, Jan Dettmers, Matthias Hemmje
Department of Multimedia and Internet Applications, FernUniversitat Hagen, Germany;
New York Institute of Technology, USA;
Department of Work and Organizational Psychology, FernUniversitat Hagen, Germany;
Zentrum fur Digitalisierung und IT (ZDI), FernUniversitat Hagen, Germany

Abstract: In this Regular Research Paper, a competency-based approach, based on the Qualifications-Based Learning Model (QBLM), will be presented, making it possible to store the acquired competencies in a Competency and Qualification Profile (CQP) during a learning unit. Before the learning unit can be started, it must be checked whether a learner has all the required Competencies for the learning unit. Therefore, an approach is shown to how the QBLM-based CQPs can be implemented, and the data from the CQP can be used to check course prerequisites. Finally, this paper presents the relevant State of the Art, the conceptual modeling, the relevant implementations, and evaluations. Finally, the paper concludes with a summary and the remaining challenges.

Empirical Study: Implementing Flex Class during Pandemic

Hyung Jae Chang, Byungkwan Jung
Department of Computer Science, Troy University, Troy, Alabama, USA

Abstract: Through the pandemic due to COVID-19, we have been experiencing many pedagogical changes in higher education. Inevitable transitions from face-to-face classes to online classes distressed many stakeholders in higher education, such as institution leaders, professors and students. Despite of the challenges in transition, Troy University adopted Flex (also called Hyflex) class format which delivers classes in 3 different ways; in-person, live lectures provided via MS-Teams synchronously and recorded lectures provided via LMS (Learning Management System) asynchronously. To observe the impact of the Flex class, we collected student participation data in 3 Computer Science Flex classes. The result depicts that many students showed interests in synchronous lectures having live interactions with instructors and other students.

Citizen Perspectives on Necessary Safeguards to the Use of AI by Law Enforcement Agencies

Yasmine Ezzeddine, Petra Saskia Bayerl, Helen Gibson
CENTRIC, Sheffield Hallam University, Sheffield, UK

Abstract: In the light of modern technological advances, Artificial Intelligence (AI) is relied upon to enhance performance, increase efficiency, and maximize gains. For Law Enforcement Agencies (LEAs), it can prove valuable in optimizing evidence analysis and establishing proactive prevention measures. Nevertheless, citizens raise legitimate concerns around privacy invasions, biases, inequalities, and inaccurate decisions. This study explores the views of 111 citizens across eight countries towards AI use by police through inter-views and integrates societal concerns along with propositions of safeguards from negative effects of AI use by LEAs in the context of cybercrime and terrorism.

A Study on Commerce Channel Characteristics and Strategies: Focus on Domestic Department Stores in Korea

Chan Uk Jung, Hyeon Jo, Won Jun Lee
SKK Business School, SungKyunKwan University, Seoul, Republic of Korea;
RealSecu Inc., Busan, Republic of Korea

Abstract: With the rapid development of digital technology, the way consumers shop continues to change. As the platforms for online commerce diversify with internet, mobile, and more, commerce channel strategy has emerged as a very important fact. In addition, due to the recent proliferation of mobile usage, customers increasingly tend to purchase via mobile channels while experiencing physical products in offline stores. This phenomenon requires traditional offline retailers to consider integrating online and mobile channels with offline stores as a commerce channel strategy. Thus, the success of traditional offline retailers depends on how well they manage their commerce channel strategy and how they implement channel services. This study analyzed commerce channel characteristics and strategies of the top domestic department stores in Korea: Lotte, Hyundai, and Sinsegae. Based on the literature, three types of commerce channels were chosen to build this analysis: single-, multi-, and cross-channel. Data on omni-channel services of domestic department stores were collected to aid in the analysis of the commerce channel strategy as well. Finally, the aim of this study was to offer data that can contribute to the development of commerce channel strategy for the Korean retail market of large-sized marts, supermarkets, convenience stores, and the like for further study.

5GD2Chain: A Blockchain Marketplace Direct to Consumer

Sergio Luis Ribeiro, Ingrid Alves de Paiva Barbosa, Suzana Cristina Pereira de Oliveira
Security and Privacy, CPQD, Campinas, Brazil

Abstract: The marketplaces new era was created to accelerate the services delivery to customers. However, the main business model for a marketplace today is to charge a commission on the amount paid by the consumer, which creates an imbalance in the distribution of value, where the actors suffer monetarily. The objective of this work is to establish a blockchain-based decentralized marketplace that eliminates the need for a centralized entity between provider and consumer, supplying commercial transactions in a more efficient and reliable way, without the need for intermediaries. This work presents the proposal and implementation results of 5GD2Chain, which is a decentralized 5G marketplace using blockchain and the Direct-to-Consumer concept, developed within the 5G BR Platform Project scope.

Mobile Report Card GUI Design Strategy to Improve the Achievement Goal

Song-Hee Kang
Hanyang University, South Korea

Abstract: By using grade cards to improve students' achievement goals through a competency-based curriculum, a strategy was presented to capture gradual development by continuously recording students\' learning processes and visualizing quantified grades and efforts. Based on the consideration of previous studies and related literature, the conceptual definition of achievement goals was understood, and the design strategy and educational technique strategy were identified as ways to promote achievement goals. Based on related literature, UI/UX design was understood as a design strategy, especially the effect of providing learning content through a web platform in user experience design, and the requirements to be considered in user experience design for students were discussed. In addition, the importance of learning outcomes of process-based evaluation as an educational technique strategy and the role of process-based evaluation from a pedagogical perspective were identified. Based on the value of the user experience, a prototype of the mobile grade card application was proposed as a mobile grade card design strategy and content research.

A Systematic Review on Mobile Application Accessibility for Low Vision Learners

Sobihatun Nur Abdul Salam, Nur Amirah, Nurulnadwan
School of Multimedia Technology and Communication, Universiti Utara Malaysia, Malaysia;
Pusat Internet Kg Tala, Kampung Tala, Jeneri, Sik, Kedah, Malaysia;
Faculty of Business and Management, Universiti Teknologi MARA (UiTM) Terengganu, Malaysia

Abstract: This study reviews an empirical studies of mobile application accessibility for low vision (LV) learners. The main aim of this review is to assist researchers developing a reference in setting a future research in mobile application for LV learners particularly in accessibility domain. A systematic review has been selected as a method for the analysis. Results from the systematic review, indicates that there is a need for the enhancement of accessibility design principles for mobile application in learning to the low vision learners. In conclusion, it is clearly revealed that the existing mobile application specifically design for accessibility design principles is still having issues and need to be addressed in future advancement towards the development of accessibility design principle of mobile application for low vision learners.

Types of Crowdsourcing and their Characteristics

Hodaka Nakanishi, Yuko Syozugawa
Teikyo University, Itabashi-ku, Tokyo, Japan;
Tokoha University, Shizuoka-Shi, Shizuoka, Japan

Abstract: There are various forms of crowdsourcing. To clarify the relationship between crowdsourcing and management strategies, we classified crowdsourcing into project type, competition type, and micro-task type based on the order workflow. In relation to existing studies, the project and competition types correspond to the aggregation type, while the micro-task type corresponds to the selection type. Differences by crowdsourcing type were analyzed with respect to indicators indicating the role of crowdsourcing (content of work, reasons for adoption, and position in the work) and indicators indicating evaluation (effectiveness of crowdsourcing, crowdsourcing usage trends, and future plans). The results showed that (1) many types of work are commonly used for all types, (2) reasons for introducing crowdsourcing match its effectiveness, and users appreciate it, (3) crowdsourcing is used according to a business strategy, and (4) many companies maintain their current use of crowdsourcing. Although crowdsourcing is an effective and highly regarded business method by users, many companies are hesitant to introduce it. To respond to recent changes in the business environment, it is necessary to reform their business processes and use the appropriate type of crowdsourcing based on the characteristics of the business.

Evaluating the Consistency of Student Performance Models for Individual Students

Enzo Novi Migliano, Kelly M. Rivera, William Frazee, Sean C. Mondesire
School of Health Science and Technology, St. Thomas University, Miami Gardens, Florida, USA;
Department of Mathematics and Statistics, Villanova University, Villanova, PA, USA;
Department of Computer Science, University of Central Florida, Orlando, Florida, USA;
School of Modeling Simulation and Training, University of Central Florida, Orlando, Florida, USA

Abstract: Online education has gained popularity in the recent decade with the demand for better remote education and learning tools. One method to improve remote education is adaptive intelligent tutoring systems. These systems present learning material to students, track student performance, predict when to intervene, and recommend material to support remediation. These systems customize the learning experience to the individual, traditionally using one student performance model for all students, implementing Markov Process-based inferences, logistic regression, or deep learning. Though these fixed approaches have performed well in specific scenarios, they limit student models to one type of static prediction technique. This paper proposes a study to examine how the effectiveness of student prediction techniques varies in performance from student to student. This work aims to validate the hypothesis that a one-size-fits-all model for performance prediction is not always optimal, and customized prediction techniques can further enhance student model accuracy.

Knowledge Transfer in e-Learning on Student Performance in Higher Education

Abeer Alsadoon, Gregory Crease, Srujal Patel
Kent Institute Australia, Sydney, Australia

Abstract: Given evidence that e-learning in the form of on-line simulation can improve higher education student performance through better enabling the transfer of learning than face to face teaching, it is important to better understand the processes that contribute to the transfer of learning through on-line simulation. This paper develops an enhanced conceptual framework of the transfer of learning through on-line simulation. Based on a review of secondary sources, this paper examines the current model of the learning transfer process, identifies addition factors and relationships that affect that model in the context of on-line simulation, and proposes an enhanced conceptual framework of the transfer of learning through on-line simulation. The enhanced framework presents a more complete understanding of the processes that affect the transfer of learning through on-line simulation, enabling educators to incorporate these processes to improve their instructional design.

A Review Study on Self-Learning Braille Systems

Amani AlSalman, Suheer Al-Hadhrami, Abdulmalik AlSalman
Special Education Department, King Saud University, Riyadh, Saudi Arabia;
Computer Science Department, King Saud University, Riyadh, Saudi Arabia

Abstract: Assistive technology is widely used by visually impaired and blind individuals, who rely mainly on the braille system for communication, reading, and writing in their daily life. Hence, the importance of developing and designing braille self-learning systems to facilitate independent learning, which will save time, effort, and cost, is a necessity. The system could be a primary tool or act as an aid tool during the learning process. This paper aims to reveal the features, deficiencies, needs, and challenges of such systems in order to improve and develop high-quality braille self-learning systems. The research review identified several braille self-learning systems; online systems, standalone device systems, computer-based systems, and smartphone-based systems. The researchers also provided recommendations for researchers and developers to achieve superior braille self-learning systems and applications to meet the targeted individuals' needs.

Supply Chain Risk Management for Dining Restaurants During the COVID-19 Pandemic

Yu Luo
Management Information System of Lerner College of Business and Economics,
University of Delaware, Newark, Delaware, USA

Abstract: The COVID-19 pandemic has severely impacted business activities around the world, especially in industries like the restaurant industry that are high-ly dependent on timely deliveries from the supply chain. Out of special con-cern for personal and family health, consumer groups have developed a preference and inclination to localize food. By analyzing consumers' willing-ness to spend in restaurants and local policies regarding COVID-19, this paper proposes a data analysis technique for different regions or environments to provide better resilience and robustness to the restaurant's supply chain, as well as to guide different regional supply chain planning and capital investment. Data at local restaurants suggest that this strategy can improve the robustness of restaurants and their upstream and downstream businesses to COVID-19.

Edge Intelligent Agent Computing Systems with Blockchain Technology in Wireless Body Area Networks: Applications and Technologies

Deok Hee Nam
College of Professional Studies, Computer and Electrical Engineering and Computer Science,
Wilberforce University, Wilberforce, Ohio, USA

Abstract: The agent-based system technology has been used frequently to improve its network connectivity and the constant miniaturization of electrical invasive/non-invasive devices to empower the development of the intelligent agent systems such as Wireless Body Area Networks (WBANs). Many technologies have proved their efficiency in supporting WBANs applications, such as remote monitoring, biofeedback, and assisted living by responding to their specific quality of service (QoS) requirements. To robust the intelligent agent systems and make the connection between WBANs, the fast-growing research fields, edge computing, and blockchain technology, are also applied in many leading applications. In this paper, a brief review of the challenges and emerging technologies for WBANs and their leading applications along with applied technologies are presented.

How Singapore's Manufacturing Small and Medium Enterprises Embrace Industry 4.0

Gopal Surianarayanan, Thomas Menkhoff
Singapore Management University, Singapore

Abstract: Based on a case study of a small manufacturing firm in Singapore, we explore how local SMEs adopt Industry 4.0 solutions. We shed light on the drivers and barriers of Industry 4.0 adoption to better understand current business dynamics, potential issues, focus areas, and initiatives to smoothen this implementation. The study is part of a wider Industry 4.0 study of key specialists and decision-makers across Government agencies, Institutes of Higher Learnings, suppliers of Industry 4.0 technology, business associations, etc. Technology push by the Government with robust funding and training support, skilled labour shortages including imported labour dependence, productivity issues and the pressure to innovate business models due to increased competition are propelling SMEs to adopt Industry 4.0. Some challenges include high investment costs, ROI concerns as well as capability and mindset issues. The paper contributes to the minimal Asian management literature about Industry 4.0 matters in Asian SMEs.

The 20th International Conference on Embedded Systems, Cyber-physical Systems, & Applications (ESCS'22: July 25-28, 2022, USA) https://american-cse.org/csce2022/conferences-ESCS https://american-cse.org/csce2022/

Fault Detection and Identification in Industrial Control Systems based on Measurement Data, a Machine Learning Approach

Sohrab Mokhtari, Kang K. Yen
Electrical & Computer Engineering Department, Florida International University, Miami, Florida, USA

Abstract: Faults and attacks in industrial control systems (ICS) are a concern for researchers due to the growing rate of malicious activities in recent years. Conventionally, model-based methods were employed for fault detection in ICSs using the mathematical model of the system. Since achieving the mathematical model is costly or even impossible, artificial intelligence using the system's data is used for fault detection purposes called learning-based strategy. This strategy proved its reliable performance in ICSs fault detection by investigating the network traffic and measurement data. The measurement intrusion detection system (MIDS) could detect abnormal behaviors in an ICS with a high accuracy rate. Although the MIDS is a reliable system for detecting abnormalities in ICSs, distinguishing the occurred fault type could increase the reliability by making it possible to compensate for the fault in a short proper time. A framework for fault detection and identification (FDI) is proposed and tested on a power system. The promising results show an accuracy of 99% for classifying the types of occurring faults in the test-bed.

Enhanced Data Security in Collaborative Intrusion Detection System

Chukwuebuka Ezelu, Ulrich Buehler
Fulda University of Applied Sciences, Germany

Abstract: Due to increasing use of internet for data exchange between companies, the borderlines between IT-Systems are constantly disappearing. New attacks scenarios become possible making single Intrusion Detection Systems (IDS) no longer a reliable defense shield. In this paper, we propose a certificateless security architecture for Collaborative IDS that incorporates private and public encryption, with entity authentication. Owing to wider range of visibility, CIDS provides a more effective and holistic protection for variety of new attack types. There still exists need to ensure secure communication between local IDSs in multiple organizations. This is achieved through bidirectional entity authentication, data encryption, and a form of homomorphic operation performed on the ciphertext for message validation. The result shows improved data and entity security and privacy.

Side Channel Attacks Based on Convolutional Neural Network with Inception Module

Feng Ni, Junnian Wang, Fanliang Hu, Wan Wang
School of Physics & Electronic Science, Hunan University of Science and Technology, Xiangtan, P. R. China

Abstract: Side channel attacks (SCAs) use the physical information leakage that accompanies the operation of encryption algorithms on encrypted hardware devices to recover the secret key. Applying deep learning algorithms to side channel analysis of hardware cryptographic chips effectively improves the efficiency and accuracy of side channel at- tack. However, most existing deep learning side channel attacks (DLSCAs) are based on traditional neural networks. In this paper, we propose an op- timised convolutional neural network side channel attack method, which applies a new network structure, Inception, to SCAs to obtain and fuse feature information in different scale perceptual fields. For the Chip- Whisperer platform encryption board running AES-128 encryption al- gorithm, DLSCAs are performed using MLP, CNN, RNN and the new network. The experimental results show that the application of the new network to DLSCAs results in fast convergence, good robustness and an increase in model accuracy of more than 11.15% compared to traditional networks.

Application of Embedded System in Solar Generating Waste Receptacle

Bassam Shaer, John Rijos-Esquilin, Nadia Kuchambi, Henry Baldwin, Jonathan Fulmer
Electrical, Mechanical & Engineering Technology Department, University of West Florida, Florida, USA

Abstract: It is evident in modern society that the purpose of most objects, often ones with single function designs, tend to be forgotten about, with their designs rarely ever being considered for improvement opportunities. In the case of this paper, we analyze the main functions of a waste receptacle, similar to ones that would be found in a park, a university campus, or any outdoor scenario, and we determine how to integrate modern solar power generation with intelligent solar tracking to maximize power output. This paper describes the processes, design methods, calculations used to develop a power generation unit from an ordinary waste receptacle. The result is a prototype of a large-capacity waste receptacle with approximately 3000W of total power provision throughout the length daylight cycle, and intelligent features such as the electro-mechanical touch-free operation of the trash flaps, and 110VAC power output capabilities of up to 500W through a DC-AC Inverter.

EverGreen - An Off-Grid Automated Vertical Farming: Design, Development, and Lessons Learned

Bahram Asiabanpour
Ingram School of Engineering, Texas State University, San Marcos, Texas, USA

Abstract: EverGreen is a multidisciplinary off-grid automated vertical farming project to conduct research and education on Food, Energy, and Water nexus, build capacity for R&D and Education, and improve the attraction and retention rates of minority students in STEM and Agriculture fields. It is a partnership program among different institutions, including Texas State University Engineering and Agriculture programs, EcoCentro, Freeman center, San Antonio College, and Palo Alto College. Under this project, a multidisciplinary off-grid vertical farming and multiple course modules were developed. Dozens of students received research training and mentorship. The dissemination of the research in his project resulted in three patents, 30 awards, and 33 refereed conference and journal articles. This article discusses the program, assessments, successes, and lessons learned.

On the Inclusion of Heterogeneous Agents in Unmanned Vehicle Swarms - A Review

Abhishek Phadke, F.Antonio Medrano, Stanislav Ustymenko, Tianxing Chu
Texas A&M University-Corpus Christi, Corpus Christi, Texas, USA;
Department of Computer science, Texas A&M University-Corpus Christi, Texas, USA;
Conrad Blucher Institute for Surveying & Science, Texas A&M University-Corpus Christi, Texas, USA

Abstract: The use of Unmanned Vehicle (UV) swarms in applications for disaster management and relief, remote sensing and surveillance, and environmental conservation has increased tenfold in recent years. The target environment for UV tasks is dynamic and can disrupt the normal workings of the swarm. Researchers have explored ideas to increase overall swarm resilience using novel methods for path planning, task allocation, network security, and overall swarm management. Additionally, including agents that differ in structure or nature from other agents of the swarm has been found to have certain advantages that increase resiliency. Such swarms are termed heterogeneous swarms. Swarm heterogeneity can be accomplished in a multitude of ways. This study is a review of current research that proposes using heterogeneous agents in a swarm to increase the overall resiliency of the swarm. Heterogeneous swarms are identified based on operational space, capability, and structure. Field applications are examined as a means to lay the groundwork for future research. Characteristics of such swarms that increase overall swarm resiliency are discussed.

An Embedded Malware Detection System Using a Support Vector Machine

Robert Oshana, Mitchell A. Thornton, Mike Caraman, Nathan Srirama
Edge Processing Business Line, NXP Semiconductors, Austin, Texas, USA;
Darwin Deason Institute for Cybersecurity, Southern Methodist University, Dallas, Texas, USA;
Edge Processing Business Line, NXP Semiconductors, Bucharest, Romania

Abstract: A prototype detection system for side channel attacks using a Support Vector Machine and processor performance counters is proposed. The performance and robustness of the approach is assessed. The system is capable of detect-ing multiple malicious exploits simultaneously. Dimensionality reduction techniques are used to improve model performance without losing accuracy. Feature selection techniques are used to determine new feature subsets. Ef-fectiveness is measured using Receiver Operating Characteristics. The ro-bustness of the detection system is measured using a Gaussian noise model and comparing the Root Mean Square Error and accuracy to the standard de-viation of noise at different noise levels. We assess the robustness of the de-tection system to CPU load using CPU stress applications and CPU utilization limits.

Digital Closed Loop IFOG Based on True-Log Amplifier

Ahmed El Ghrabawy, Mostafa Fedawy, Mohamed Saad Elmahallawy
Electronics & Communication Department, College of Engineering,
Arab Academy for Science, Technology, and Maritime Transport, Cairo, Egypt;
Mechatronics Department, Higher Technological Institute, Ramadan City, Egypt;
Center of Excellence in Nanotechnology, Arab Academy for Science, Technology and Maritime Transport, Egypt

Abstract: Response time, output steady-state error, angle random walk (ARW), bias instability (BI), and implementation simplicity are the main challenges that faces the closed loop interferometer fiber optic gyroscope (CL-IFOG). In this paper, a new design for the CL-IFOG based on true-log amplifier is proposed to improve its performance. A comparison between conventional and proposed CL-IFOG is carried out. The comparison shows an average improvement in the transient response time and steady state error by 60.3% and 61.4% respectively over wide range of rotation rates. Also, the ARW and BI are improved by 5.4% and 42.9% respectively. Moreover, the ADC and DAC resolutions are reduced by 25% and 16.7% respectively which simplify the hardware implementation.

Detecting Code Injections in Noisy Environments Through EM Signal Analysis and SVD Denoising

Ekaterina Miller, Georgios Michail Makrakis, Kurt A. Vedros, Constantinos Kolias, Craig Rieger, Daniel Barbara
Department of Computer Science, University of Idaho, Idaho Falls, Idaho, USA;
Idaho National Lab, Idaho Falls, Idaho, USA;
George Mason University, Fairfax, Virginia, USA

Abstract: The penetration of embedded devices in networks that support critical applications has rendered them a lucrative target for attackers and evildoers. However, traditional protection mechanisms may not be supported due to the memory and computational limitations of these systems. Recently, the analysis of electromagnetic (EM) emanations has gathered the interest of the research community. Thus, analogous protection systems have emerged as a viable solution e.g., for providing external, non-intrusive control-flow attestation for resource-constrained devices. Unfortunately, the majority of current work fails to account for the implications of real-life factors, predominantly the impact of environmental noise. In this work, we introduce a frame work that integrates singular value decomposition (SVD) along with outlier detection for discovering malicious modifications of embedded software even under variable conditions of noise. Our proposed framework achieves high detection accuracy i.e., above 93% AUC score for unknown attacks, even for extreme noise conditions i.e., -10 SNR. To the best of our knowledge, this is the first time this realistic limiting factor, i.e., environmental noise, is successfully addressed in the context of EM-based anomaly detection for embedded devices

Charging Scheduling of Electric Vehicles using Charge Time Priority

Myrzabek Murataliev, Hakduran Koc
College of Science and Engineering, University of Houston-Clear Lake, Houston, Texas, USA

Abstract: In this paper, we focus on the coordinated charging of electric vehicles considering the impact they would have on the electric grid infrastructure especially during peak hours. After studying Start Time Priority and Finish Time Priority algorithms, we propose a novel algorithm to solve the charging scheduling problem considering charge time priority (CTP). The CTP algorithm schedules the vehicles considering their charging time interval requests, travel time to the charging station and queuing time in addition to the actual charging time. The algorithm also equally distributes the long time-taking charging processes to available charging stations/outlets. The experimental evaluation performed using various charging scenarios clearly shows viability of the proposed algorithms that outperform the existing approaches significantly.

Initial Study of the Effectiveness by Using Coral USB Accelerator to Accelerate Goal Recognition on CanSat

Miho Akiyama, Hiroshi Ninomiya, Takuya Saito
Graduate School of Electrical and Information Engineering, Shonan Institute of Technology, Fujisawa, Japan;
Department of Information Science, Faculty of Engineering, Shonan Institute of Technology, Fujisawa, Japan;
Department of Informatics, Faculty of Informatics, Tokyo University of Information Science, Chiba, Japan

Abstract: This paper describes a method to recognize a goal by deep learning on a CanSat using the Coral USB Accelerator, which is faster than the conventional method. In the CanSat contest, the goal is recognized by the camera image, and how to lead the robot to the goal at a distance of 0m is a competition. We have previously achieved goal recognition with high accuracy by using deep learning and won several prizes at ARLISS 2019. However, due to size constraints, CanSat can only be equipped with computers with low computing power such as Raspberry Pi Zero. Therefore, we investigated how much faster it could be made by using the Coral USB Accelerator.

Compressing the Compressor; Reducing VHDL Generated FFTs Using Memory Table Isolation and Compression

Michael Kapralos, Ethan Shafer
Department of Electrical Engineering & Computer Science, United States Military Academy,
West Point, New York, USA;
University of Connecticut, Connecticut, USA

Abstract: FPGA designs for numerical applications are often optimized for speed and accuracy, with the amount of available hardware treated like a secondary objective. Balancing speed, accuracy, and hardware uti- lization is a difficult proposition, especially when multiple instances of the same design, or multiple different designs are used on the same chip. In this paper, we demonstrate an automated process for identifying and minimizing memory tables within VHDL code which then can be used to balance these resource trade-offs. Using established minimization methods, traditionally used for programmable logic devices (PLDs), we convert individual output bits into minimal logic circuits, while allowing the parallelization of the minimization process improving the speed. After introducing the basics of the algorithm, we demonstrate the utility of this process with direct comparisons to examples from a previous work. It is demonstrated that the amount of CLB Look-Up Tables (LUTs) used in a design can be reduced by as much as 33% in large designs, with some types of LUTs being reduced by as much as 54%. The process is also quick, and simple, requiring a few minutes more, at most, than the regular synthesis process.

The 18th International Conference on Foundations of Computer Science (FCS'22: July 25-28, 2022, USA) https://american-cse.org/csce2022/conferences-FCS https://american-cse.org/csce2022/

A Method of Counting to Simplify Solving Recurrence Relations

Daniel Spiegel, Danielle S. Hope
Kutztown University of Pennsylvania, Pennsylvania , USA;
T-Mobile, Atlanta, Georgia, USA

Abstract: One of the more daunting topics in math disciplines as well as computer science is deriving a closed-form solution for a sequence or a recurrence that represents a sequence and proving it by induction. Common methods taught to find a closed-form include expanding a recurrence and identifying patterns. Similarly, inductive proof is often presented with an imprecise process. In this work a method of finding a closed-form solution by counting expansions represented by a sequence will be presented and a process for demystifying the proof will be demonstrated.

On Arguments and Argument Forms in Discrete Computational Structures

Ahmed Tarek, Ahmed Alveed, Ahmed Farhan
Engineering, Physical & Computer Sciences (EPCS), SET Div, Montgomery College, Rockville, Maryland, USA;
The Wharton School, University of Pennsylvania (UPenn), Philadelphia, Pennsylvania, USA;
The Johns Hopkins University, School of Medicine, Baltimore, Maryland, USA

Abstract: Arguments, and argument forms are of paramount importance in Discrete Computational Structures, in Legal Analysis and in underlying theory pertaining to Software Engineering to infer valid conclusion based upon the presented facts, premises, and realities. This paper explores arguments and argument forms, and the arguments are differentiated from non-arguments. Classification has been made depending upon the nature of arguments or argument forms that they represent. Examples and instances are considered for clarification. Emerged as an important logical tool in computing, several significant results are also presented. General conclusion is inferred based upon the structures of arguments, argument forms, and their potential applications. There are not many scientific research papers that have explored this avenue of research. From that perspective, this paper has significant contributions to Computing.

Partial Mirror Matrices and Symmetric Boolean Functions

Peter M. Maurer
Department of Computer Science, Baylor University, Waco, Texas, USA

Abstract: Exotic symmetries are those that go beyond simple permutations and permutation matrices. These types of symmetries are generated by matrix groups that are isomorphic to the symmetric group of degree n, , but not conjugate to the group of all permutation matrices, . It has already been shown that mirror matrices can be used to generate one type of exotic symmetry. Here we show two new types of exotic symmetry that are generated by partially mirrored matrices. We also characterize the orbit structure of all currently known exotic symmetries.

Two New Exotic Symmetry Classes

Peter M. Maurer
Department of Computer Science, Baylor University, Waco, Texas, USA

Abstract: Matrix-based symmetry of Boolean functions has been shown to be a useful tool for analyzing and simulating Boolean circuits. Until recently the useful groups have been the set of all permutation matrices, its conjugates, and the sub-groups thereof. However, there are other matrix groups, not conjugate to the group of permutation matrices, that could be useful in describing various types of symmetry. We call this type of symmetry exotic symmetry. To use the exotic matrix groups, it is necessary to have a clear understanding of how they are formed. Using the tools of mirror matrices and partially mirrored matrices, the formation of several such groups can be described. This paper introduces two new types of exotic symmetry that are not conjugate to the group of permutation matrices or to the exotic symmetries discovered so far.

An Improved Sudoku Solver

Adiba Afif Suha Binta Wadud, M. Abdullah-Al-Wadud
Bangladesh International School, English Section, Riyadh, Saudi Arabia;
Department of Software Engineering, College of CIS, King Saud University, Riyadh, Saudi Arabia

Abstract: Sudoku is a popular puzzle game played by people of all ages. Consequently, many methods are developed to solve these puzzles on computer. However, these methods have deficiencies, especially in terms of the time taken to solve the puzzles. In this paper, we propose a method which is more efficient than the methods used before. This uses a few additional data structures and updated algorithmsto make the method faster at solving various Sudoku puzzles.

Approximate Sorting and Sequence Analysis

Aditya Narasimhan, Sridhar Radhakrishnan, C. R. Subramanian
School of Computer Science, University of Oklahoma, Norman, Oklahoma, USA;
The Institute of Mathematical Sciences, HBNI, Taramani, Chennai, India

Abstract: Sorting algorithms are a fundamental problem in computer science and can demand a lot of resources when asked to sort a large set of n elements (in the order of billions). Known is that these n elements can be completely sorted using O(n log n) comparisons. However, to reduce the resources used, we take two perspectives in this paper look at the underlying patterns and sequences in the input; and to leverage the output quality. For the later approach, we put a constraint on the number of comparisons D > 0, given as an input to the sorting algorithm that performs the sorting in O(Dn) comparisons. To measure the ”sortedness” of the output with the given constraint to the algorithm, there are quality measures defined. For the first approach, we provide and investigate algorithms that analyze the patterns in the input order. This knowledge of patterns are defined along with specific metrics and corresponding promised upper bounds, that performs approximate sorting in O(nD) comparisons.

An Efficient Algorithm for Mining Most Popular Packages

Yangjun Chen, Bobin Chen
Department of Applied Computer Science, University of Winnipeg, Canada

Abstract: Given a set of items, and a set of user preferences, we investigate the problem of designing a most *popular package* (or say, a *pattern*), i.e., a subset from the items that maximizes the number of satisfied users. It is a typical problem of data mining. In this paper, we address this issue and propose an efficient algorithm for solving the problem based on a graph structure, called a p*-graph, used to represent the preference of a user, by which a lot of useless checks can be avoided. The time complexity of the algorithm is bounded by O(n^2m^3), where m is the number of items (or say, attributes) and n is the number of user preferences.

On the Application of P2SH Bitcoin Scripts

Yong Zhang
Department of Computer Science & Information Technology, Kutztown University of Pennsylvania, USA

Abstract: We study the application of Bitcoin scripts. We show that P2SH Bitcoin Scripts can be constructed based on NP-complete problems Subset Sum and Hamiltonian Cycle. Such scripts lead to more flexible transaction types and have applications in areas such as crowdsourcing computations.

The 18th International Conference on Frontiers in Education: Computer Science & Computer Engineering (FECS'22: July 25-28, 2022, USA) https://american-cse.org/csce2022/conferences-FECS https://american-cse.org/csce2022/

Educating Digital Natives

Matthew N. O. Sadiku, Uwakwe C. Chukwu, Abayomi Ajayi-Majebi, Sarhan M. Musa
Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, Texas, USA;
Department of Engineering Technology, South Carolina State University, Orangeburg, South Carolina, USA;
Department of Manufacturing Engineering, Central State University, Wilberforce, Ohio, USA

Abstract: The phenomenon of digital technologies has produced a distinctively new generation defined by digital media: the digital natives. Digital natives are those who were born into the digital age or after 1980. This generation of students are growing up in a digital world and educators must employ new approaches to make learning real and relevant to them. There is a need for our educational systems to listen to the voices of the digital natives or face a future with disgruntled, disconnected learners. Research on digital natives is important for developing 21st-century education. This paper presents some ways educators can understand digital natives and change their teaching styles accordingly. It also intends to encourage students, teachers, school administrators, and families to collaborate in developing learning environments best suited for 21st-century learners.

Multiple Fault Diagnosis in Digital Circuits using Critical Path Tracing and Enhanced Deduction Algorithm

Mohamed Mahmoud, Hussain Al-Asaad
Department of Electrical & Computer Engineering, University of California Davis, California, USA

Abstract: This paper has developed an effect-cause analysis technique for fault diagnosis in digital circuits. The main algorithm of our technique is based on the Enhanced Deduction Algorithm, which processes the real response of the CUT to the applied test T to deduce the values of the internal lines. An experimental version of the algorithm has been implemented in C++. The code takes about 7592 lines. The internal values are determined based on the logic values under the permanent stuck-fault model. Using backtracking strategy guarantees that the actual values are covered by at least one solution, or no solution is found.

Undergraduate Research Experience in Internet of Things and its Impact on Academic Experience

Hassan Salmani, Mohsen Mosleh
Department of Electrical Engineering & Computer Science, Howard University, Washington, DC, USA;
Department of Mechanical Engineering, Howard University, Washington, DC, USA

Abstract: Research in the field of Internet of Things (IoT) at the early years of undergraduate studies might be seen as challenging owing to the lack of sufficient training and knowledge. To discuss this challenge, this paper presents the assessment of an undergraduate research experience (URE) program. The participants in the URE were either freshmen or sophomores with a Grade Point Average (GPA) of less than 3.0- populations susceptible to attrition from engineering. The research activities span over five months in three different years for different cohorts. The results of knowledge gains, effects on academic performance, and student activities are discussed. Initial results show enhanced retention and increased students' GPA.

Do Students Perceive the Benefits of e-Learning Higher When They Feel Mentally Challenged?

Hyeon Jo
Department of Strategic Planning, RealSecu, Busan, South Korea

Abstract: Since the emergence of COVID-19, university students have been educated online. Risk perception of COVID-19 and mental discomfort caused by isolation may affect students' perception of the benefits of e-learning. This research aims to identify the determinants of e-learning benefits. A total of 160 samples were collected by surveying university students in South Korea. This study conducted partial least squares–structural equation modeling (PLS-SEM) method to empirically analyze the data. The findings showed that benefits are affected by affective risk perception and cabin fever syndrome.

Project Method: An Experience Report

Gabriel Bradford, Abby Ortego, Bonnie Achee, Ghassan Alkadi
Southeastern Louisiana University, Hammond, Louisiana, USA

Abstract: As part of the Industry Connect Initiative in the Department of Computer Science at Southeastern Louisiana University, students are required to complete multiple real-world-ready classes [1,6]. This paper serves as an experience report of a student project for Software Engineering I, the first of these classes. Project Method is a website that allows a student to organize their academic responsibilities into one space. Various technologies, which will be discussed in further detail within this paper, were implemented in order to assemble the front-end and back-end. This class presented students with a real-world scenario to develop soft skills, such as interacting with a client, working as a team on a project, and experience with Agile software development.

Industry Connect Initiative - On the Importance of an Industry Advisory Board in an Undergraduate CS/IT Program, An Experience Report

Bonnie Achee, Ghassan Alkadi, Matthew McNulty, John Burris
Southeastern Louisiana University, Hammond, Louisiana, USA

Abstract: Southeastern Louisiana University.s Department of Computer Science established an Industry Connect Initiative that strives to provide students with real-world ready-skills, best practice awareness and a relevant curriculum. Industry Connect Initiative is a four-pronged approach that includes an Industry Advisory Board (IAB), a distinguished lecture series, an internship program and industry-inspired curriculum integrations. The Industry Advisory Board (IAB) is made up of key personnel from businesses in the area. Several of the businesses represented were founded by alumni. The IAB serves as the linchpin for the Industry Connect Initiative. The IAB both informs and directs all other areas of the initiative. This paper details the current organization and involvement of the Department of Computer Science IAB at Southeastern Louisiana University with a brief overview of its formation. Recommendations and lessons learned are presented.

Industry Connect Initiative: The Distinguished Lecturer Series

Bonnie Achee, Matthew McNulty, Ghassan Alkadi
Southeastern Louisiana University, Hammond, Louisiana, USA

Abstract: The Distinguished Lecturer Series (DLS) was founded in the Fall semester of 2020 as part of the Industry Connect Initiative at Southeastern Louisiana University. The purpose of DLS is to bring together industry leaders in a variety of areas of computer science to provide informative talks to the students and faculty of the Computer Science Department and to provide an informal venue for interactions to build relationships that will serve to be mutually beneficial for all. DLS provides both students and faculty with an opportunity to explore topics and best practices in a variety of areas, expanding their breadth of knowledge and therefore the discipline as a whole. This poster paper will discuss the Distinguished Lecturer Series and demonstrate how this prong of the Industry Connect Initiative serves that mission.

MOC 6 - Using the EMDS-5 Dataset to Classify Microorganisms for Education

Max Cole, Allen Mire, Ghassan Alkadi, Bonnie Achee
Southeastern Louisiana University, Hammond, Louisiana, USA

Abstract: This project explores the creation of a multi-disciplinary capstone project, MOC 6, a web application used to identify submitted microorganisms. The motivation behind this project was exploration into a topic that appears to be mostly unexplored along with the ability to utilize the recently released EMDS-5 dataset in our exploration. The ability to classify microorganisms from images would be beneficial to amateurs, hobbyists, and educators who may lack the technical knowledge of the field required to identify them on their own.

Visualizing General Education Curriculum Assessment Data

Cesur Dagli, Molly Hall, Bethany Bodo
Office of Analytics and Institutional Effectiveness, Virginia Tech, Virginia, USA

Abstract: This poster outlines the assessment and visualization process of Virginia Tech's new Pathways to General Education curriculum, which was implemented for incoming freshmen starting in the fall of 2018. The Institutional Effectiveness unit in the Office of Analytics and Institutional Effectiveness at Virginia Tech facilitates Pathways assessment with the goal of providing data that can be used to inform and improve the Pathways program to enhance student learning at Virginia Tech. Individual courses and faculty also utilize their own data to make improvements on a more granular level. The general education program has 39 student learning outcomes and 500+ courses. Data visualizations completed to date are also presented.

Collaborative Learning to Teach Set Theory to Engineer Students

Jorge Juarez, Maria del Rayo Guevara-Villa, Anabel Sanchez Sanchez, Raquel Diaz Hernandez, Leopoldo Altamirano Robles
Instituto Nacional de Astrofisica, Optica y Electronica, Mexico;
Universidad Politecnica de Puebla, San Mateo, Mexico

Abstract: We designed a new method to teach engineer and computer science students set theory. This method is based on collaborative learning where the students form a team, and each member has a different role. Each group is divided into three subteams called the ”checker,” the ”modeler,” and the ”builder” subteams. The first step is to build a relation between the basic concepts of the set theory like union, intersection, and complement with the digital logic basic concepts like disjunction, conjunction, and negation. Later, the students analyze a set theory problem making a digital logic equivalent. The checkers carry on this step. Then the modelers design a truth table with all possibilities. The builders construct the logic circuit equivalent using TTL gates in a breadboard. They use all information obtained for the three subteams (checker, modeler, and builder). The whole team can now model problems relating set theory with digital logic. This method was used to teach set theory to computer system engineering and biotechnology engineering students, obtaining a better knowledge of this mathematics area.

A Telepresence Robot for a Better Orientation at University

Laurent Gallon, Timothee Duron, Marine Rigal
Universite de Pau et des Pays de l'Adour, E2S UPPA, LIUPPA, Mont-de-Marsan, France;
Universite de Pau et des Pays de l'Adour, E2S UPPA, SCUIO-IP, Pau, France

Abstract: This article presents the first results of the use of telepresence robots to allow high school students absent from our University, to discover our courses in telepresence, and assist them in their academic pathway. After presenting the project, we detail the process built to use the robot. Then we give the results of a survey conducted among the first pilots regarding the immersive activity itself, and the impact of the use of the telepresence robot.

A Top-Down Approach for Teaching Programming Language – Teaching Python via an Open-Source ERP Framework

Ning Chen
Department of Computer Science, California State University, Fullerton, California, USA

Abstract: In a bottom-up fashion, the traditional pedagogy of teaching programming language asks learners to learn the parts of a programming language at the bottom level and integrate them later at the top level. As a comparison, we propose a top-down teaching approach in which learners first play the role of a user of an open-source ERP framework (Odoo, in this study), subsequently, the role of requirement engineer, tester, and developer. This study shows the following benefits: a natural way of a human learning sequence – curiosity, playing, observation, imitation, and innovation; a more inclusive self-learning, open-sourced environment for all types of learners; same as the bottom-up style, the top-down approach also complies with the ACM-IEEE CS curricula 2013.

A Systematic Way to Enhance Student Engagement in a Machine Learning Module

Jie Cheng, Mehdi Zargham
Department of Computer and Information Technology, Miami University, USA;
Department of Computer Science, University of Dayton, Ohio, USA

Abstract: This paper represents a new way of teaching a machine-learning course. Students starts with a set of assignments to implement simple neural networks in Python and progress toward assignments to perform object detection within an image. The assignments will complement the theories of learning derived in the class. Students will use Jupyter Notebook platform to code the lab assignments and pro-jects. Each project is built on the prior projects and encourages students to achieve better results by implementing class concepts. Students are able to visualize their data to gain more about dependency between the attributes. Students will also use Google TensorFlow to do major practical applications. Students are sometime surprised with their own results and try to use this technology to create new apps. They write codes for creating a apps and are excited to do changes as necessary. They make every effort to do a good project while having fun and learning new materials.

Virtual ABET Accreditation: Benefits and Challenges

Mudasser F. Wyne, Shakil Akhtar
College of Professional Studies, National University, USA;
CS/IT Department, Clayton State University, USA

Abstract: ABET accredited programs are widely recognized nationally and internationally. To date, over 4,000 programs at 850 colleges and universities in 41 countries have received ABET accreditation. Over 175,000 students graduate from ABET-accredited programs each year and join millions of graduates who have received degrees from ABET-accredited programs since 1932. ABET accreditation requires a systematic assessment and evaluation of student outcomes on a regular basis. Presentation of evidence for such assessment of student outcomes and individual course learning outcomes during and before accreditation visit is required for accreditation. ABET has adapted the accreditation visit practices in view of the recent changing circumstances due to Covid-19. In this paper we share challenges as well as opportunities presented for virtual accreditation visits. Some researchers have drawn parallels between virtual accreditation visits and distance teaching and learning, online education, and online consultations. This paper will highlight these connections and will further identify some of the challenges in conducting and preparing for virtual accreditation visit. Examples from the practical experience acquired during national and international accreditation visits are presented. The paper also reports on experience while being the host of one of such virtual accreditation visits. The finding of this research is expected to bridge the gap between the academic institution and accreditation agencies. In addition, the paper also promotes virtual accreditation visits leading to cost and time saving, while reducing the pollution and environmental harm due to excessive travel.

A Cross-Discipline Capstone Project Experience

Li Yang
Department of Computing and Mathematics, University of West Georgia, Carrollton, Georgia, USA

Abstract: Computer Science capstone course is a senior-level course in which students integrate and apply what they have learned in their under-graduate curriculum to a substantially large-scale software develop-ment project. Authentic real-world projects motivate students to be more engaged and give them the opportunity to interact with real customers. In this paper, we describe our experience in a real-world cross-discipline capstone project that involved joint effort and com-munication of students and faculty in CS and Education. We present the project background and the structure as well as the lessons that we've learned and the benefits and impact of such a project.

Capstone Experiences in Developing Augmented Reality Tables for Community Organizations

H. Keith Edwards, Michael R. Peterson, Francis Cristobal
Department of Computer Science, University of Hawaii at Hilo, Hawaii, USA

Abstract: This paper examines two Senior Capstone experiences developed as augmented reality tables over the past two years. Both projects were public facing efforts that required working implementations. The first project was at an astronomy center and focused on interactions between land use and ecological aspects of the Big Island while the second project focused more on historical sites. Both projects leveraged brownfield development and existing code bases to allow for student success in spite of the impacts of the COVID-19 pandemic.

Team-Based Online Learning in Multidisciplinary Research and Instruction

Matthias K. Gobbert, Jianwu Wang
Department of Mathematics and Statistics, University of Maryland, Baltimore County, Maryland, USA;
Department of Information Systems, University of Maryland, Baltimore County, Maryland, USA

Abstract: There is a critical nationwide shortage of IT professionals as well as of scientists and engineers with high-performance computing (HPC) and big data related advanced computing skills. Simultaneously, the technology is growing in complexity and sophistication, which has led to the use of multi-disciplinary teams with members from a broad range of home domains everywhere in industry, government, and academia. Moreover, a lot of the vital team collaborations take will place virtually using a variety of software platforms now and in the future. We report here on experiences with preparing undergraduate and graduate students for these career opportunities in several contexts, from regular semester classes, an undergraduate summer research program, to an advanced graduate student CyberTraining program. All these programs are conducted fully online and leveraged concepts of flipped classrooms, recorded lectures, team-based and active learning, regular oral presentations, and more to ensure student engagement and lasting learning.

The SPORT-C Intervention: An Integration of Sports, Case-Based Pedagogy and Systems Thinking Learning

Jeffrey Basoah, William Scherer
University of Virginia, Charlottesville, Virginia, USA

Abstract: The STEM field is unrepresentative of the population it serves. Due to a lack of cultural relevance in STEM courses, there is a dissociation between the lived experience of students from underrepresented racial groups (URG) and STEM course material. The SPORT-C intervention is a framework that combines sports, systems thinking learning, and a case-based pedagogy into an activity that can be used in any STEM course. A pilot study was conducted to determine the viability of the SPORT-C intervention in a classroom setting and determine if it was worth further investigating and if any impact differed by racial identity. The findings from this study implicate that the SPORT-C intervention has an impact on the motivation levels of students to participate in STEM courses.

A Study on the Utilization of Mobile Education System to Improve the Achievement Goal

Song-Hee Kang
Hanyang University, South Korea

Abstract: The recently revised curriculum presents process-oriented evaluation as a change in the evaluation paradigm. Process-oriented evaluation is an evaluation that comprehensively evaluates the learner's learning process with various evaluation methods and evaluation contents to help the learner's overall growth. With the advent of new evaluation terms, it is necessary to find out the definition and evaluation direction. Therefore, this study presented a strategy to capture gradual development by continuously recording students' learning processes and visualizing quantified grades and efforts by using report cards to improve students' achievement goals through a competency-based curriculum. First, the essential purpose of the grade card is to identify the current location, recognize individual strengths and weaknesses of students, and help them find solutions to them. It is necessary to maximize usability and increase concentration with a concise, accessible interface design through appropriate functions and practical information that meet goals and intentions, so that applications can be used naturally like habits. Second, in the past, grade information was digitized and listed in descriptive form, making it difficult to understand and less readable, making it inconvenient for users to use. Therefore, it aims to provide motivation for achievement goals by visualizing grade data by collecting satisfaction surveys and student requirements of existing apps, as well as delivering grade information through emoticons with dynamic motion graphics technology. Third, it induces awakening by providing goal scores of career-wanted universities to visually check their own target college entrance information to promote their willingness to practice achievement goals and provide entrance examination information such as past competition rates so that grades can be managed efficiently with one app. This study aims to bring about meaningful changes to improve students' achievement goals by redesigning grade cards centered on students. It is expected that this study will not only solve educational problems in the future, but also be used as basic data as an important educational communication tool actively used in the school field in the composition of student-centered mobile education services.

Improving the Performance of BSCS and BSIT Courses by Applying an Instant Messenger Tool in an Online Learning Environment

Yanzhen Qu, Howard Evans, Noura Abbas, Mazen Alkhatib
School of Computer Science, Colorado Technical University, Colorado Springs, Colorado, USA

Abstract: One of most distinct learning experiences for students enrolled in an online only degree program is that they conduct most academic activities through an Online Learning Environment (OLE). Students, in the same class, do not always share the same lecture hours together like students in a classroom of a ground campus. Students mostly listen to archived lecture recordings to learn the new course materials. The engagement between students and instructors is more “on-demand” and “individualized”. This study examines the utilization of an instant messenger tool with a set of rules or criteria can result in enhanced engagement for students learning in an OLE. The focus of our study is to evaluate a set of Bachelor of Science in Computer Science (BSCS) and Bachelor of Science in Information Technology (BSIT) courses’ completion rates and students’ engagement in their course assignments based on a set of adopted faculty/student weekly messaging rules/criteria. Our study has demonstrated that our adopted instant messenger tool strategies can be effective in achieving higher course completion rates and higher levels of engagement in discussion board assignments in an OLE.

Segue Toolkit for Data Science

Courtney S. Lamar, Lethia S. Jackson
Department of Cyber-Physical Systems, Clark Atlanta University, Atlanta, Georgia, USA;
Department of Technology and Security, Bowie State University, Bowie, Maryland, USA

Abstract: The Segue Toolkit for Data Science provides "plug-and-play" learning activities for faculty to integrate into foundational programming courses. This Toolkit consists of three components: (1) Python Primer; (2) Statistics Review and Refresh; and (3) Python-based Data Science Activities. The supplemental learning activities are based on Process-Oriented Guided Inquiry Learning (POGIL), a learner-centered instructional approach where students learn in teams of 3-4 students with specially designed instructional materials that follow A.E. Lawson's 3-phase "learning cycle." Having been successfully piloted in two (2) courses at Clark Atlanta University and Bowie State University, preliminary results show increased student confidence levels in understanding data science and proficiency in demonstrating the data science life cycle.

An Integrated Project/coach-Based Teaching Approach in an Undergraduate Software Engineering Course: Preliminary Observations

Robert S. Ellinger, Marc Boumedine
University of the Virgin Islands, St. Thomas, Virgin Islands, USA

Abstract: A one-semester Software Engineering course was offered by a faculty and an Enterprise Architect (guest lecturer and coach) with more than 40 years of experience in various industries. The regular project-based learning (PBL) approach was redesigned by integrating PBL and coaching-based learning approaches. The coach introduced a patented rapid application development (RAD) process, particularly suited for web applications. Students learned from first-hand anecdotes the meanings and the impacts of "dos" and the "don'ts" as well as the "Fast, good, cheap, pick two" expressions with real-life examples. They quickly connected those principles with the course material usually presented in lectures. Students integrated the RAD principles in their term project. Integrating PBL and the Coaching approaches fills significant gaps identified by software industries, specifically during requirement analysis.

BATTLE 2022: Preparation of Future Cyber Technologists

Clarence Ray, Jesse Bemley
Bowie State University, Bowie, Maryland, USA

Abstract: This research presents the initial two years of Battle and a proposed third year which builds on the successes and failures of the previous years. The project represents a continuing effort by the Bowie State University Department of Technology and Security to develop and implement programs to assist average students in enhancing their academic and social credentials. In addition to technical skills a heavy emphasis is placed on further development of soft skills to develop a well-rounded student. They became adept at navigating corporate cultures and environments. The twenty-five students recruited were sophomores, juniors, and seniors. Other facets were further development of communication skills and technical documentation. Over the 5 week period, the technical and social growth was very evident.

Critical Factors for Recruiting Women in Computer Science in the US Virgin Islands and Beyond

Michael T. Francois, Rebecca Hoffart, Marc Boumedine
University of the Virgin Islands, St. Thomas, Virgin Islands, USA

Abstract: This paper reviews studies to examine important factors for women in order to pursue a degree computing related. The findings will hopefully inform us and guide our strategies for recruiting women in computer science at the University of the U.S. Virgin Islands (UVI) and at similar academic institutions. In 2020, 64% of students enrolled at the University of the Virgin Islands were women. However, only 20% are majoring are majoring Computer Science. According to the Bureau of Labor Statistics women only earn 18% of the computer science bachelor degrees in the United States. This first study in the US Territory will help determining what factors contribute the most in attracting or detracting women Virgin Islanders from chosen a career in computer science. This preliminary work synthetizes the findings from existing research and will examine whether additional factors are important and specific for Virgin Islanders women.

Development of a Synergistic Education Framework for Tech-Oriented Startups and Scaleups Leaders

Dong H. Lee, Jong H. Lee
Center for Technology Startup Education Team, Research Institute of Convergence, DGIST, Daegu, Korea

Abstract: We develop a synergistic education framework to bridge startup and scaleup program and introduce the innovative technology venture education platform that can systematically share tech-based startups and scaleups leaders. It gives a good platform to create new knowledges related to technology venture startups and obtain the ability to various management skills and R&D strategies for each stage of corporate growth.

Teaching Parallel Computing to a Lower-Level Programming Course

Xiaoyuan Suo
Department of Math and Computer Science, Webster University, USA

Abstract: Parallel and Distributed Computing (PDC) topics are traditionally found in up-per-level computer science courses. Due to the complexity nature of the content, PDC is rarely introduced to the lower levels, i.e., freshmen or sophomore level. The purpose of this work is to study the feasibility in introducing PDC content to a lower-level freshman CS programming course. This article then shows the result and lessons learned after introducing Parallel and Distributed Computing (PDC) thinking strategies using unplugged activities to my first and second-year undergraduate students, who are enrolled in computer programming courses.

Team Projects: Cross-pollinating Environmental Science and Technology (ESTEC) Using Project-Based Learning

Velma C. Latson, Deidre Gibson
Department of Technology & Security, Bowie State University, Maryland;
Hampton University, Virginia, USA

Abstract: Engaging students in different domains is a challenge. This project brings together undergraduate students from two renowned HBCU communities, Bowie State University and Hampton University, to cross-pollinate computing technology and environmental science. This paper describes the success and challenges experienced over three years of engaging students in a project-based learning environment.

Empathic Design in Computer Engineering: Accessibility in Sport for Deaf Curling Athletes

Marcia R. Friesen, Robert D. McLeod
Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada

Abstract: This work presents three years of development and reflection for an introductory course in Digital Systems Design in an undergraduate computer engineering program at a Canadian university. The course was moved from a theory-driven course to a team project-based course built on a framework of design for accessibility and equity principles for deaf curling athletes. This is a report from its third year of offering, following on a 2020 paper that presented the course's first offering. Students are loosely guided through a modified design cycle, from conceptual design to a functional prototype with ergonomic considerations of a physical prototype. The course includes a strong emphasis on allied topics for modeling, knowledge transfer, including business development, technology marketing, intellectual property protection, as well as moving toward commercialization. At time of writing, the course has been offered twice through to completion and is in its third offering. It is an effective model to learn and appreciate design, although not without teething pains, as an instance of pedagogy evolution.

An Automated SQL Query Grading System Using an Attention-Based Convolutional Neural Network

Donald R. Schwartz, Pablo Rivas
School of Computer Science and Mathematics, Marist College, Poughkeepsie, New York, USA;
School of Engineering & Computer Science, Department of Computer Science, Baylor University, Waco, Texas, USA

Abstract: Grading SQL queries can be a time-consuming, tedious and challenging task, especially as the number of student submissions increases. Several systems have been introduced in an attempt to mitigate these challenges, but those systems have their own limitations. This paper describes our novel approach to automating the process of grading SQL queries. Unlike previous approaches, we employ a unique convolutional neural network architecture that employs a parameter-sharing approach for different machine learning tasks that enables the architecture to induce different knowledge representations of the data to increase its potential for understanding SQL statements.

Strengthening the Management Information Systems Course Syllabus

Nadia Khalique, Saqib Ali, Abdul Khalique Shaikh
Department of Information Systems, Sultan Qaboos University, Oman

Abstract: The rising demand for Information System (IS) workers in numerous domains is indicated by the broad information systems job market and pervasive usage of technology in various disciplines. Since the inception of IS in mid of 20th century, the discipline of information systems has seen enormous expansion in breadth and complexity demonstrating an emergent need of modernizing IS courses offered at higher educational institutes. This paper attempts to present a model for review and revision of the Introduction to Management Information Systems course offered at an undergraduate program based on the Computing Curricula (CC2020) guidelines, an endeavor supported by the Association for Computing Machinery (ACM) and the IEEE Computer Society. CC2020 is a vision for the future of computing education that includes a detailed study that contrasts pedagogical recommendations and contextualizes them within the larger computing education environment. In the course review and revision process, a competency-based model has been developed that can be used by any undergraduate degree program in the international higher education that plans to redesign its Information Systems core courses. The key learnings suggest the inclusion of emerging technologies within the course along with associated pedagogies like visual aids and case studies, to meet the needs and demands of today's information systems.

For Process-Oriented Evaluation - A Comparative Study on the Education Evaluation System

Song-Hee Kang
Hanyang University, Republic of Korea

Abstract: In the 2015 revised curriculum in Korea, Process-Oriented Evaluation is presented as a change in the evaluation paradigm. Process-Oriented Evaluation is an evaluation that comprehensively evaluates learners' learning processes with various evaluation methods and evaluation contents to help learners grow overall. Learning, in which its own evaluation area is not a visible part such as technology and knowledge, but an internal motive, that is, attitude and growth, are steadily revealed in the long term, is required. An appropriate evaluation should be planned and implemented. The results show that students' colorful competencies are revealed in the same context and showing the process over a long period of time needs to be presented in the changed curriculum. The form of class could be predicted through the analysis of evaluation methods of individual countries. Rather than serving as a channel for the transmission of knowledge, classes are conducted as a structure in which students themselves can learn through various activities and feedback suitable for the characteristics of each subject [1,2]. In the case of Korea, it is difficult to expect learning development through teachers' feedback on evaluation. In the end, this evaluation method reduces the competence of teachers. The fact that students do not get good grades in multiple-choice and short-answer tests conducted in schools can lead not only to failure to go to school, but also to deteriorate the evaluation of guidance teachers. Evaluation for selection is also necessary, but for students, evaluation should be a tool for future development. This study aims to bring about meaningful changes to improve students' achievement goals by newly designing grade cards centered on students. It is expected that this study will be used as basic data as an important educational communication tool actively used in the school field in the composition of student-centered educational services as well as solving future educational problems.

Integrating Mobile App Development in Computing Education: Our Experience

Weidong Liao, Osman Guzide
Department of Computer Science, Mathematics and Engineering, College of STEM,
Shepherd University, Shepherdstown, West Virginia, USA

Abstract: In this paper, we discuss our efforts to alleviate the recruitment and retention challenges in four-year STEM degree programs by integrating Mobile App Development into our computing curricula. We will describe our practice of using project-based learning in capstone classes, and revamping an advanced level Web programming course to cover mobile app development so as to provide a career-oriented, team learning environment with mobile app development projects to attract and retain students, especially underrepresented students in computing programs, and prepare them for more career opportunities. Our practice also emphasizes design creativity, application development, writing and presentation, teamwork, and collaboration, by utilizing mobile app development as a focal point.

Detection and Classification of Dropout Behavior

T. Duron, L. Gallon, P. Aniorte
Universite de Pau et des Pays de l'Adour, LIUPPA, Mont de Marsan, France

Abstract: This article presents the first results of our work aimed at detecting students' risk of dropping out and dropping out during block-based computer programming lessons in the classroom. We use information related to the use of the mouse and the interface of the programming software to detect behaviors characteristic of dropout. The paper describes the methodology chosen to build the rules for detecting these behaviors. The results obtained on a first case study are also given. These results were obtained within the framework of the PERSEVERONS project1, which aims to study the perseverance of students induced by the use of a digital object.

Obsolescence of Computer Science Education

Joel Weymouth, Ramesh K. Karne
Computer and Information Sciences, Towson University, Towson, Maryland, USA

Abstract: Since universities first established undergraduate and postgraduate Computer Science programs, there have been studies and papers on what a program should include in a Computer Science (CS) discipline. There have also been articles concerning what research areas should be part of the CS curriculum. A search of the IEEE database shows articles written as early as 1961 on Computer Science Education (CSE). A cursory analysis suggests that the CSE discipline has been a relevant subject of research and discussion for over sixty years. This paper examines the IEEE/ACM Computing Curriculum recommendations over the years. The 1968 Recommendation by ACM/IEEE is the baseline since Computer Science Education was just beginning to be a recognized curriculum. Then entries from 1991, 2001, 2008, 2013, and 2020 to test a hypothesis that Computer Science has evolved from purely theoretical to primary application with some emphasis on trade rather than innovation. In short, computer science has been made obsolete and that is unfortunate. The question is whether the motivation was societal, business, or governmental needs that drove this evolution in CSE. Analysis of each report will show these answers

The 18th International Conference on Grid, Cloud, & Cluster Computing (GCC'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-GCC
https://american-cse.org/csce2022/

Cloud-native RStudio on Kubernetes for Hopsworks

Gibson Chikafa, Sina Sheikholeslami, Salman Niazi, Jim Dowling, Vladimir Vlassov
Hopsworks AB, Stockholm, Sweden;
KTH Royal Institute of Technology, Stockholm, Sweden

Abstract: In order to fully benefit from cloud computing, services are designed following the "multi-tenant" architectural model, aimed at maximizing resource sharing among users. However, multi-tenancy introduces challenges of security, performance isolation, scaling, and customization. RStudio server is an open-source Integrated Development Environment (IDE) accessible over a web browser for R programming language. We present the design and implementation of a multi-user distributed system on Hopsworks, a data-intensive AI platform, following the multi-tenant model that provides RStudio as Software as a Service (SaaS). We use the most popular cloud-native technologies, Docker and Kubernetes, to solve the problems of performance isolation, security, and scaling present in a multi-tenant environment. We further enable secure data sharing in RStudio server instances to provide data privacy and allow collaboration among RStudio users. We integrate our system with Apache Spark, which can scale and handle Big Data processing workloads. Also, we provide a UI where users can provide custom configurations and have full control of their RStudio server instances. Our system was tested on a Google Cloud Platform cluster with four worker nodes, each with 30GB of RAM. The tests on this cluster showed that 44 RStudio servers, each with 2GB of RAM, can be run concurrently. Our system can scale out to potentially support hundreds of concurrently running RStudio servers by adding more resources (CPUs and RAM) to the cluster or system.

The 8th International Conference on Health Informatics & Medical Systems (HIMS'22: July 25-28, 2022, USA) https://www.himscsce.org/hims22 https://american-cse.org/csce2022/

Recent Applications of Deep Learning in Health Informatics: A Review

Farzan Shenavarmasouleh, Farid Ghareh Mohammadi, Khaled M. Rasheed, Hamid R. Arabnia
School of Computing, University of Georgia, Georgia, USA;
Institute of Artificial Intelligence, University of Georgia, Georgia, USA

Abstract: Deep learning (DL) has bestowed us with powerful analyzing tools and techniques, which researchers can use and apply in various fields of study. Health informatics generates the most amount of data of any discipline in today's era and can therefore benefit exceedingly from DL. Extracting features and finding complex patterns from a huge amount of raw data and transforming them into knowledge is a challenging task. Besides, various DL architectures have been proposed by researchers throughout the years to tackle different tasks. This paper provides a review of DL models and their broad application in bioinformatics and healthcare categorized by their architecture. In addition, we also study some of the key challenges that still exist and can show up while conducting DL research.

Design and Developing an Accurate and Low-Cost Automated Reader System for Rapid HIV Testing Kits

Ali Elkateeb
Electrical and Computer Engineering, University of Michigan, Michigan, USA

Abstract: N/A

Introduction of Heart Rate Measurement by Facial Video to Hugging-Type System

Wang Chao, Takehiko Ogawa, Ho Yihsin, Jun Hasegawa, Naoki Oshima
Mechanical and Electronic Systems Engineering Course, Takushoku University, Tokyo, Japan;
Department of Electronics and Computer Systems, Takushoku University, Tokyo, Japan;
Department of Design, Takushoku University, Tokyo, Japan

Abstract: A hugging type measurement system can quickly measure vital signs while the patient remains calm. In this system, various sensors are attached to the sphere to measure respiration rate, heart rate, body temperature, etc. A photoelectric pulse-wave sensor is a common method for measuring heart rate with a fingertip. In recent years, a pulse-wave measurement method based on facial images has been proposed and implemented. Herein we proposed a method that uses facial images captured by a camera for heart rate measurement in a hugging system. A camera was attached to the sphere, and an experiment was conducted to measure the heart rate in a posture hugging the sphere.

Why Does AI Fail in Healthcare?

Aaron Lai, Thomas Ming
Health Industry Advisory, PwC Advisory, San Francisco, California, USA;
California Northstate University Medical School, Elk Grove, California, USA;
HKUSPACE Community College, Hong Kong

Abstract: AI has made significant inroad in many industries and we wonder why healthcare applications are less impressive. Even equipped with latest technology, vast amounts of data, significant funding, big names tech companies, and a determined drive, there are still more high-profile failures than success stories. This article argues that a successful AI system in clinical setting requires three things: an accurate result for obvious reason, an Explanatory AI to show its reasoning to the clinician, and a Conversational AI to allow interrogation between a clinician and the system so that the clinician can form true collaboration with the machine. Knowledge acquisition for an AI system is then discussed with reference to clinical setting and reasoning.

Application of Dimensional Reduction in Artificial Neural Networks to Improve Emergency Department Triage During Chemical Mass Casualty Incidents

Nicholas D. Boltin, Joan M. Culley, Homayoun Valafar
Department of Computer Science & Engineering, University of South Carolina, Columbia, South Carolina, USA;
College of Nursing, University of South Carolina, Columbia, South Carolina, USA

Abstract: Chemical Mass Casualty Incidents (MCI) place a heavy burden on hospital staff and resources. Machine Learning (ML) tools can provide efficient decision support to caregivers. However, ML models require large volumes of data for the most accurate results, which is typically not feasible in the chaotic nature of a chemical MCI. This study examines the application of four statistical dimension reduction techniques: Random Selection, Covariance/Variance, Pearson's Linear Correlation, and Principle Component Analysis to reduce a dataset of 311 hazardous chemicals and 79 related signs and symptoms (SSx). An Artificial Neural Network pipeline was developed to create comparative models. Results show that the number of signs and symptoms needed to determine a chemical culprit can be reduced to nearly 40 SSx without losing significant model accuracy. Evidence also suggests that the application of dimension reduction methods can improve ANN model performance accuracy.

Human Activity Recognition on Time Series Accelerometer Sensor Data using LSTM Recurrent Neural Networks

Chrisogonas O. Odhiambo, Sanjoy Saha, Corby K. Martin, Homayoun Valafar
Computer Science & Engineering, The University of South Carolina, Columbia, South Carolina, USA;
Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, Louisiana, USA

Abstract: The use of sensors available through smart devices has pervaded everyday life in several applications including human activity monitoring, healthcare, and social networks. In this study, we focus on the use of smartwatch accelerometer sensors to recognize eating activity. More specifically, we collected sensor data from 10 participants while consuming pizza. Using this information, and other comparable data available for similar events such as smoking and medication-taking, and dissimilar activities of jogging, we developed a LSTM-ANN architecture that has demonstrated 90% success in identifying individual bites compared to a puff, medication-taking or jogging activities.

The Effects of Physical Activity Levels on Sleep in Shift-Work Nurses During the COVID-19 Pandemic

Sang-Ho Lee, Yeongmi Ha, Sanghun Yun, Oh-Seok Kwon, Dong-Ha Lee, Jinung An
Division of Intelligent Robot, DGIST, Daegu, South Korea;
College of Nursing, Gyeongsang National University, Jinju, South Korea;
Division of Electronics and Information System, DGIST, Daegu, South Korea

Abstract: The aim of this study was to compare the sleep characteristics and problems according to physical activity levels in shift-work nurses during the corona pandemic. To this end, 57 female shift work nurses were recruited for this study. Physical activity and sleep quantity were measured using a wearable device, and sleep quality and problems were evaluated with the Korean versions of the Pittsburgh Sleep Quality Index (PSQI-K), Insomnia Severity Index (ISIK) and Epworth Sleepiness Scale (KESS). Data were analyzed with Student's t-test using SPSS version 19.0. The participants were placed into two groups: the low physical activity group was sedentary for over 639.5 minutes per day, and the high physical activity group was sedentary for under 639.5 minutes per day. The amount of physical activities influenced the following sleep variables after the night shift: total sleep time ($p = .017$), sleep time ($p = .017$), wake after sleep onset ($p = .040$), deep sleep ratio ($p = .015$), light sleep time ($p = .017$) and rem sleep time ($p = .026$). Therefore, maintaining an adequate physical activity level has a positive impact on sleep for female shift-work nurses.

Development of Tracking System for Welfare Facility Using iBeacon

Kenichi Takahashi, Takao Kawamura
Graduate School of Sustainability Science, Cross-informatics Research Center, Tottori University, Japan

Abstract: In a welfare facility for intellectually impaired persons, the staff must recognize when and where intellectually impaired persons have caused concerns, such as wandering at night. Subsequently, a staff member is required to visit the area where the concern has occurred to prevent residents from accidents. Thus, welfare facilities entreat a system to support the discovery of such concerns. As the first step in the support system, we developed a tracking system using iBeacon. We analyzed the questionnaires from the staff in welfare facilities and designed and implemented functions based on the analysis results.

Engineering Puzzles: Designing a Data Collection Device for Cognitive Assessment

R. J. Guncheon, N. M. J. Hince, D. B. Stewart, L. Zhan, Marcia R. Friesen
Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada

Abstract: The work designed and investigated the use of serious games as a complementary tool to assess and detect mild cognitive impairment (MCI). The paper details the design and development of a fully functional prototype of a physical maze game. The maze is instrumented to track a person's trajectory through a maze, including any hand tremors during their play. The maze is a tool to collect empirical data to aid in MCI assessment, "a syndrome defined as cognitive decline greater than expected but one that does not interfere notably with activities of daily life" [1]. Current methods of MCI diagnosis are time consuming and intensive, which can impact the timeliness of successful interventions. If data collected from the maze can differentiate between persons with and without MCI, the maze game may be used to support complementary diagnosis methods. The instrumentation of the maze includes sensors for positioning to track a player's moves through the maze, hand tremor detection, processors for handling and storing data, and a web page to display a user's results.

Self-Monitoring Diabetic Lungs Pre and Post Physical Activity Through Deep Learning/AI Techniques

Hoda Salsabili, Miad Faezipour
School of Engineering Technology, Electrical & Computer Engineering Technology, Digital/Biomedical Embedded Systems and Technology (D-BEST) Lab, Purdue University, West Lafayette, Indiana, USA

Abstract: Diabetes affects an estimated 34.2 million people (10.5 percent of the United States population) and is the seventh leading cause of death. Diabetes complications often lead to hospitalization and impaired quality of life due to vascular damage. Physical activity is highly recommended to prevent and treat diabetes complications. Recently, self-monitoring techniques have received attention to improve and track the recovery of diabetes people during and after physical activity. Given the damage to the lung vascularity in diabetes, there is a lack of accessible and easy-to-use techniques to inform changes in the lung volume and breathing patterns. In the literature, machine/deep learning and Artificial Intelligence (AI) have been demonstrated to be promising techniques for tracking the changes in breathing patterns using a variety of modalities (signal/images). This paper proposes an AI-guided analysis of the respiratory pattern from breathing sounds for diabetes. We designed a Convolutional Neural Network (CNN) to take breathing sound signals and convert them into images of the frequency spectrogram to identify the breathing phases. AI-based regression is then used to estimate the lung volume as a function of the energy of the extracted breathing sound phases. The ideas allow the creation of computational patient-specific lung profiles that can trace the effects of physical activity on the lung recovery of diabetes.

Healthcare Data Analytics: Application to Prostate Cancer

Alfadil Ahmed Hamdan, Fathe Jeribi, M-Tahar Kechadi, Mohammed A. Fakhruddin, Mostafa M. Mehanawi
Department of Computer Science, College of CS & IT, Jazan University, Jazan, Saudi Arabia;
Department of Information Technology and Security, College of CS & IT, Jazan University, Jazan, Saudi Arabia;
School of Computer Science, University College Dublin, Dublin, Ireland;
Department of Computer Networks, College of CS & IT, Jazan University, Jazan, Saudi Arabia

Abstract: In recent years, healthcare data analysis is becoming one of the most promising research areas. Healthcare datasets include data in various types such as clinical data, Omics data, and sensor data, etc. Clinical data includes electronic health records that store patient records collected during pathological periods and continuous treatment. Omics data is one of the high dimensional data comprising genome, transcriptome and proteome data types. While Sensor data is collected from various wearable and wireless sensors. It is very hard to handle this raw data with simple statistical models. Machine learning techniques have emerged as an important tool for large data analysis. There are many types of algorithms of machine learning, such as supervised, unsupervised and reinforcement. In this research paper we show the ability of machine learning techniques to analyze healthcare data. We used prostate cancer dataset collected from National cancer Institute (NCI) as a case study. After data processing we applied different classification techniques such as Artificial Neural networks (ANN), Random forest (RF), Decision Tree (J48), Logistic Regression (LR) and Naïve Bayes. A k-fold cross-validation methodology was used in model building, evauation and comparison. The results showed that ANN and RF are the most accurate predictor (with a test set accuracy of 100%) for this domain, followed by decision tree 99.82%, Logistic regression 98.76% and Naive Bayes 96.64%.

A Fusion of Transformers for Automatic Interpretation of Biomedical Images in Scientific Literature

Md Mahmudur Rahman, Oyebisi F. Layode
Computer Science Department, Morgan State University, Baltimore, Maryland, USA

Abstract: The manual interpretation and generation of knowledge from medical images is not only time-consuming and error prone, but also impractical. Hence, systems that can automatically map visual content present in the images to concise textual representations is a necessity to assist and speed up the diagnosis process of medical professionals. Automatic natural language interpretation of images is a challenging task which combines Machine Learning (ML), Computer Vision (CV), and Natural Language Processing (NLP) fields. Currently, the medical imaging field is witnessing a growing interest for Transformers that can capture global context compared to Convolutional Neural Networks (CNNs) with local receptive fields. Inspired from this transition, this work aims to interpret and summarize insights gained from images in a large collection image appeared in biomedical literature in the form of caption prediction and concept detection by utilizing and fusing both text and vision-based transformer in Sequenceto-sequence learning. Experimental evaluation and result analysis in a large collection of radiology images in terms of F1 score for concept detection and BLUE score for caption prediction showed promising result and also justify our fusion approach with the best results.

Clinical Decision Support System an Easy Access to Streamline Basic Healthcare Services in Regional QLD

Saad Masood Butt, Carmen Reaiche, Tracey Mahony, Stephen Boyle
James Cook University, Australia

Abstract: As a result of the shortage of health practitioners in Regional Queensland and other conditions such as limited nursing staff, access to diagnostics equipment due to natural disasters can be a limiting factor to affordable, basic healthcare. Given these limiting factors there is a growing need to embrace alternative technologies such as decision support system (DSS) that will enable timely and efficient access to basic healthcare. To evaluate the impact of DSS in improving the efficiency and increasing access of basic healthcare services as well as to stimulate the adoption of such systems in healthcare services this study employed a descriptive design which involved a review of recent peer-reviewed literature published on the topic of DSS in healthcare. The study established that health DSS can help in dealing with the shortage of Health Practitioner, eliminate the geo-graphical barriers that hinder access to healthcare and improve the timely access to affordable and quality basic healthcare. The research indicated significant potential in enhancing access to affordable and quality healthcare. However, challenges due to high initial costs of equipment and staff training as well as lack of adequate legal support limit the possible adoption of such systems. Therefore, to mitigate the slow uptake of DSS, there is a need to have a good will from relevant policy developers so that they can create a favorable framework to accelerate the development and uptake of DSS.

Machine Learning Applied to the Analysis of Information on Infectious Diseases Type Covid-19 in the Mlabnarino Project

Colombia Sixto Enrique Campana Bastidas, Ana Isabel Vallejo, Rigoberto Roberto Rosero
Universidad Nacional Abierta y a Distancia - UNAD, Colombia;
Hospital Universitario Departamental de Narino - HUDN, Colombia

Abstract: The pandemic generated by COVID-19 put all the world health systems on alert, making knowledge to treat these types of conditions a priority, where human survival is at risk. But thanks to technologies such as Artificial Intelligence and its different Machine Learning models, it has been shown that these tools can help in the analysis and prediction of behaviors that allow people to mitigate the risk of fatal disasters due to infectious diseases such as COVID-19. Based on that, this paper presents the advances that have been implemented in the MlabNariño project, where, based on the strengthening of a molecular laboratory to carry out tests for infectious diseases, machine learning models have also been incorporated to classify information and predict possible situations and complications that would occur with patients according to their symptoms. The applied methodology is descriptive, working on the data of the patients using the molecular laboratory, with their prior informed consent. The results are preliminary, not yet conclusive, but they promise important applications to help people mitigate the risks of the new global threat that, apart from wars, now appears to be pandemics.

Hybrid Matching Methods for Treatment Program Evaluation: A Case Study

Nafis Neehal, Georgios Mavroudeas, Malik Magdon-Ismail, Jason Kuruzovich, Kristin, P. Bennett
Rensselaer Polytechnic Institute, Troy, New York, USA

Abstract: We study a type-2 diabetes (T2D) health management program (HMP) using causal methods for treatment effect estimation on electronic health records. We use matching and survival analyses to assess T2D onset and acute care usage (emergency room or inpatient visits). To account for bias and healthcare usage changes due to the COVID-19 pandemic, we developed a hybrid matching approach that first identifies the set of potential controls based on time and other critical features and then applies matching methods. We compare results across seven state-of-the-art methods including expert-informed approaches. We find that HMP potentially improved subject health by more rapidly identifying patients with undiagnosed T2D at enrollment, allowing for timely treatment. After the initial two months, no significant differences are observed in time to T2D onset. We also found that HMP patients were less likely to seek acute care indicating improved health outcomes. We highlight practical challenges in observational health studies.

Blockchain-Enabled Virtual Healthcare Applications using Mobile Edge Computing

Janet Light-Thompson, Colin LeMesurier
Department of Computer Science, University of New Brunswick, Saint John, NB, Canada

Abstract: The development of Internet of Things has enabled smart healthcare for patient monitoring, health management, etc. An environment where smart healthcare systems can thrive is facilitated by the technology that enables the fifth generation of mobile cellular 5G systems. Mobile edge computing is a key technology of 5G networks, providing optimization of mobile resources, processing of data prior to sending it to the cloud, and reducing transmission latency. However, issues arise with vulnerabilities in security and privacy. Blockchain and smart contract technology provide a solution to security vulnerabilities while providing additional benefits to a smart healthcare environment. This paper proposes a framework which combines mobile edge computing with blockchain technology to provide mobile access to health data while providing benefits of data security, privacy, and integrity. A prototype is developed which utilizes the Truffle framework to create a decentralized application using the Ethereum blockchain for data collection and storage of patient medical information.

Quality Criteria Analysis in Process Models Discovered for Breast Cancer Therapy Marie

Marie Bommersheim, Patrick Philipp
Fraunhofer Institute of Optronics, System Technologies & Image Exploitation IOSB Karlsruhe, Germany

Abstract: A possible application for AI-based clinical decision support systems is the complex therapy in advanced breast cancer. Thereby, quality analysis of AI models can contribute to the trustworthiness of the systems. In this context, our analysis of therapy processes has shown that a model that considers all patients strongly generalizes and hardly allows precise conclusions. Dividing patients into homogeneous groups, grouping individual drug administrations, and using a heuristic approach has resulted in models in which generalization and precision are more balanced. This balance, as well as increased simplicity of the models, may contribute to the trustworthiness of a decision support system in this area.

Validation of a Neighborhood Spot Evaluation Method for Walking Route Recommendation

Katsuyuki Yamauchi, Panote Siriaraya, Da Li, Felix Dollack, Yukiko Kawai, Shinsuke Nakajima
Kyoto Sangyo University, Japan;
Kyoto Institute of Technology, Japan;
Fukuoka University, Japan

Abstract: In recent years, the population of people walking for the purpose of dieting and maintaining physical fitness is on the rise. Even though walking has significant health benefits, many users find it difficult to go for a walk on a regular basis. To support walking as an exercise, we are working on the development of a support system that enables users to enjoy walking in a positive manner. In this paper, we calculate ratings for gourmet spots based on word-of-mouth analysis and describe the construction of a walking route recommendation system based on these values. In addition, we conducted an experiment with participants to verify the validity of using this system to find spots in the vicinity.

Four Things People Should Know about Migraines

Mohammad S. Parsa, Lukasz Golab
University of Waterloo, Waterloo, Ontario, Canada

Abstract: Migraine literacy among the public is known to be low, and this lack of understanding has a negative impact on migraineurs' quality of life. To understand this impact, we use text mining methods to study migraine discussion on the Reddit social media platform. We summarize the findings in the form of four things people should know about chronic migraines: it is a serious disease that affects people of all ages, it can be triggered by many different factors, it affects women more than men, and it can get worse in combination with the COVID-19 virus.

Mobile Applications as Complementary to Traditional ADR Reporting: Encouraging Citizenship Participation and Performing Longitudinal Studies

Marzia Di Filippo, Alessandro Avellone, Michael Belingheri, Maria Emilia Paladino,
Michele Augusto Riva, Antonella Zambon, Dario Pescini
Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy;
School of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy

Abstract: In response to the last advances of mobile apps as complementary tools to traditional pharmacovigilance surveillance systems, we proposed a new mobile app called vaxEffect@UniMiB and we assessed its potentialities in fostering citizenship engagement to an adverse drug reactions reporting campaign. Specifically, we evaluated the reporting rate in an academic cohort vaccinated for COVID-19, together with its capability to monitor the same subject over time. While guaranteeing full privacy of the submitted data, vaxEffect@UniMiB turned out the first attempt in performing longitudinal studies to analyse per individual time series in terms of the registered adverse events after each vaccine dose.

The 24th International Conference on Artificial Intelligence (ICAI'22: July 25-28, 2022, USA) https://american-cse.org/csce2022/conferences-ICAI https://american-cse.org/csce2022/

An Automated Deduction of the Halting Problem

Jack K. Horner
Independent Researcher, Lawrence, Kansas, USA

Abstract: The Halting Problem – that there is no effective procedure for determining whether a program running on a Turing machine halts -- is the best-known limit to comprehensive verification of software. Here I describe what appears to be a novel proof of that Problem, based on Burkholder's 1987 first-order axiomatization.

An Automated Equational Logic Derivation of "I avoid kangaroos" in Lewis Carroll's Symbolic Logic

Jack K. Horner
Independent Researcher, Lawrence, Kansas, USA

Abstract: Lewis Carroll's introductory symbolic logic textbook, Symbolic Logic, contains exercises that require the reader to discover "the" consequent of a given set of sentences. Here, I first "manually" solve Exercise 60 in that textbook, then use the automated equational logic deduction system contained in Mathematica to corroborate the solution. The automated corroboration proof appears to be novel.

Addressing Usable-Security Challenges for NLP Systems

Yasser M. Hausawi
Deputy Director General for Shared Services & Associate Professor of IT & Cybersecurity,
Center for Digital Transformation & Information, Institute of Public Administration, Riyadh, Saudi Arabia

Abstract: National Language Processing (NLP) is a widely used approach in many computing and humanities areas. There is considerable number of research work covering all aspects of NLP, but very few consider and look at the NLP systems from the usability-security angle. This contribution is part of a series of continuous research work that looks at NLP systems from that angle to help evaluating and enhancing those NLP systems in both usability and security. A model was previously authored and published to help people better understand the nature of the relationship between NLP systems on the one hand and usability-security from the other. Also, a matrix was previously authored and published to evaluate and enhance NLP systems based on usability and security points of view. In this paper, both of the previously proposed model and matrix were evaluated though a long-term paired-sample t test comparative study of 30 projects used by 924 participants between Spring 2015 and Fall 2019. The results showed that the proposed work has clear impact on achieving usable-security on NLP systems.

Machine Learning Land Cover and Land Use Classification of 4-Band Satellite Imagery

Lorelei Turner, Torrey Wagner, Paul Auclair, Brent Langhals
Graduate School of Engineering & Management, Air Force Institute of Technology,
Wright-Patterson AFB, Ohio, USA;
The Perduco Group (a LinQuest Company), Beavercreek, Ohio, USA

Abstract: Land-cover and land-use classification generates categories of terrestrial features, such as water or trees, which can be used to track how land is used. This work applies classical, ensemble and neural network machine learning algorithms to a multi-spectral remote sensing dataset containing 405,000 28x28 pixel image patches in 4 electromagnetic frequency bands. For each algorithm, model metrics and prediction execution time were evaluated, resulting in two families of models; fast and precise. The prediction time for an 81,000-patch group of predictions was <1 s for the fast models, and >5 s for the precise models, and there was not a significant change in prediction time when a GPU was used. Logistic regression was the best fast model, with >91% accuracy & f1 on the holdout dataset. While this model struggled with the trees and grassland classes, it provided inferences that the green wavelength band and the perimeter of the blue wavelength were important. All precise models had better performance than prior work, and the best precise model resulted from an 8-hyperparameter simultaneous neural network optimization. This model possessed >97% accuracy & f1 on the holdout dataset, and >97% accuracy & f1 on each of the 6 land categories. For both best-of-family models, there was little overfitting. Finally, the model was validated by predicting the land categories of a sample 159 MB multispectral image and visually verifying correct predictions. Using these classification models to automatically monitor land-cover and land-use classification is a promising approach for tracking changes over time and potentially reducing analyst workload.

Effect of Connection State and Transport/Application Protocol on the Machine Learning Outlier Detection of Network Intrusions

George Yuchi, Torrey Wagner, Paul Auclair, Brent Langhals
Graduate School of Engineering& Management, Air Force Institute of Technology,
Wright-Patterson AFB, Ohio, USA;
The Perduco Group (a LinQuest Company), Beavercreek, Ohio, USA

Abstract: The majority of cyber infiltration & exfiltration intrusions leave a network footprint, and due to the multi-faceted nature of detecting network intrusions, they are often difficult to detect. In this work a Zeek-processed PCAP dataset containing the metadata of 36,667 network packets was modeled with several machine learning algorithms to classify normal vs. anomalous network activity. Principal component analysis with a 10% contamination factor was used to identify anomalous behavior. Models were created using recursive feature elimination on logistic regression and XGBClassifier algorithms, and also using Bayesian and bandit optimization of neural network hyperparameters. These models were trained on a dataset with numeric features, and also with the addition of categorical variables related to connection state, transport state and application protocol. The XGBClassifier algorithm generated near-perfect models with an f1 metric on the train/test dataset of 0.994 or better for both datasets. The mean accuracy of the best model on each dataset was 99.9%, which compares favorably to prior machine learning-assisted NIDS work that possessed a mean accuracy of 96.8%. The addition of state and protocol variables gave a slight improvement to modeling, and the XGBClassifier algorithm showed the best model performance. Notably, when recursive feature elimination was applied to best model, performance was sustained with the removal of transport layer protocol information.

Effect of Trigonometric Transformations on the Machine Learning Prediction and Quality Control of Air Temperature

Andrea Fenoglio, Torrey Wagner, Paul Auclair, Brent Langhals
Graduate School of Engineering & Management, Air Force Institute of Technology,
Wright-Patterson AFB, Ohio, USA;
The Perduco Group (a LinQuest Company), Beavercreek, Ohio, USA

Abstract: Conducting effective quality control of weather observations in real time is vital to the 14th Weather Squadron's mission of providing authoritative climate data. This study explored automated quality control of weather observations by applying multiple machine learning techniques to 43,487 surface weather observations from 5 years of data at a single location. Temperature predictors were evaluated using recursive feature elimination on linear regression and XGBoost algorithms, as well as using a neural network hyperparameter sweep. Modeling was repeated after calculating trigonometric transforms of temporal variables to give the models insight into the diurnal heating cycle of the Earth. All models developed in this study demonstrated better performance than a trivial model and demonstrated an acceptable <5% level of overfitting between the training and holdout datasets. The best model resulted from a 77-network hyperparameter sweep on the transformed dataset. That neural network model contained three hidden layers of 80 neurons, possessed a 0.67 C mean absolute error, and 1% overfitting as measured on the training and holdout datasets. Models on the transformed dataset exhibited a modest 1-2% performance improvement over the untransformed dataset, indicating that providing a temperature-prediction model insight into the diurnal heating cycle of the Earth may facilitate other improvements. Given this greater understanding of weather quality control that results from machine learning modeling, there is a potential to limit the time that weather analysts require to maintain accuracy of historical weather records.

Increasing UAS Sensor Fusion Autonomy with Spatial Voting Enhanced Spatio-Temporal Memory

James Crowder, John Carbone
Cobham Advanced Electronic Solutions (CAES), Advanced Technology and Engineering (AT&E), Colorado Springs, Colorado, USA;
Forcepoint Global Governments and Critical Infrastructure, Richardson, Texas, USA

Abstract: Unmanned vehicles must operate in unstructured environments that are inherently unpredictable and dynamical. An autonomous UAS must have some degree of intelligence to undertake tasks without direct and continuous human involvement, especially in unknown environments. This includes storing, correlating, and retrieving spatial and temporal tags and information associated with sensor readings. Critical then, are not only cognitive spatial-temporal context architecture, but artificial split-second decision making. A Spatio-Temporal Database Memory (STDM) system's initial design capable of storing and correlating complex spatio-temporal information and effective-ness example is presented. Finally, advanced Information Theoretic Methods (ITM)-Spatial Voting is adapted for enabling rapid UAS self-learning, spatial self-awareness and for increasing autonomous multi-modal sensor fusion and analyst situational context.

A Novel CNN for Low-Grade Gliomas Classification On MRI

Jun Liu, Wenbin Zhang, Geng Yuan, Xue Lin, Xiaolin Xu, Yanzhi Wang
Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA;
Heinz College, Carnegie Mellon University, Pittsburgh, PA, USA;
Department of Electrical & Computer Engineering, College of Engineering, Northeastern University, Boston, Massachusetts, USA

Abstract: Codeletion of chromosomal arms 1p/19q has been connected to a good response to treatment in low-grade gliomas (LGG) in several studies. The ability to predict 1p19q status is critical for treatment planning and patient follow-up. The goal of this research is to develop a noninvasive approach based on MR images using our novel convolutional neural networks. Although public networks such as VggNet, GoogleNet and other well-known public networks can effectively diagnose brain cancers using transfer learning, but the model contains a huge number of components that are unrelated to medical image. As a result, the diagnostic results are unreliable by transfer learning model. In order to address the issue of trustworthiness, we build the model from the bottom up, rather than relying on transfer learning. Our network structure flexibly uses a deep convolution stack mixed with a dropout and dense operation, which reduced overfitting and enhanced performance. We also augmented the given dataset. The Gaussian noise is introduced during the model training. We use three–fold cross-validation to train the best selection model. Our proposed model is compared to MobileNetV2, InceptionResNetV2 and VGG16 that have been fine-tuned through transfer learning. Our model achieves better results than these models. When classifying images between 1p/19q codeletion and not codeleted, the proposed architecture achieved an F1-score of 96.37%, Precision 97.46%, Recall 96.34% in the test set.

Reservoir Computing and Echo State Networks: Heuristics for Architecture, Parameters, and Initialization

B. Weborg, G. Serpen, R. A. M. Arnest
Electrical Engineering & Computer Science Department, University of Toledo, Toledo, Ohio, USA;
National Aeronautics and Space Administration (NASA), USA

Abstract: Echo state network, a recurrent neural network, avoids many of the shortcomings of recurrent networks. It offers a relatively easy training process with prediction performance that is superior to simple recurrent networks. As there are many parameters requiring initial value selections, it is likely to lead to an extended trial and error process for the novice user and therefore poses a steep learning curve. Among the most important design decisions is the specification of the reservoir which is randomly configured. Intuitively, a random reservoir generation is not an optimal solution, having negative impacts on the training and prediction performance, thus it is imperative to set parameters in such a way to guarantee better random reservoir instantiations. Thus, the formulation of heuristics through empirical study, to assist new users in developing correct architecture, by setting parameters and their initial values, is the main objective of this research. We consider four benchmark problems for an extensive simulation study to identify the factors that affect the prediction and training performance and provide recommendations to choose the architecture, parameters, and initialization for specific problem domains as well as across the collection of the problem domains.

Multiscale Graph Neural Networks for Protein Residue Contact Map Prediction

Kuang Liu, Rajiv K. Kalia, Xinlian Liu, Aiichiro Nakano, Ken-Ichi Nomura,
Priya Vashishta, Rafael Zamora-Resendiz
Collaboratory for Advanced Computing & Simulations, University of Southern California, Los Angeles, USA;
Department of Computer Science, University of Southern California, Los Angeles, USA;
Department of Physics & Astronomy, University of Southern California, Los Angeles, USA;
Department of Chemical Engineering & Materials Science, University of Southern California, Los Angeles, USA;
Department of Quantitative & Computational Biology, University of Southern California, Los Angeles, USA;
Department of Computer Science, Hood College, Frederick, Maryland, USA;
Lawrence Berkeley National Laboratory, Berkeley, California, USA

Abstract: Machine learning (ML) is revolutionizing protein structural analysis, including an important subproblem of predicting protein residue contact maps, i.e., which amino-acid residues are in close spatial proximity given the amino-acid sequence of a protein. Despite recent progresses in ML-based protein contact prediction, predicting contacts with a wide range of distances (commonly classified into short-, medium- and long-range contacts) remains a challenge. Here, we propose a multiscale graph neural network (GNN) based approach taking a cue from multiscale physics simulations, in which a standard pipeline involving a recurrent neural network (RNN) is augmented with three GNNs to refine predictive capability for short-, medium- and long-range residue contacts, respectively. Test results on the ProteinNet dataset show improved accuracy for contacts of all ranges using the proposed multiscale RNN+GNN approach over the conventional approach, including the most challenging case of long-range contact prediction.

Effects of Technology Acceptance Factors on Satisfaction with Virtual Personal Assistants

Hyeon Jo
Department of Strategic Planning, RealSecu, Busan, Republic of Korea

Abstract: Artificial intelligence (AI) has evolved steadily and the market is growing rapidly. AI helps consumers' lives in the form of a virtual personal assistant (VPA). The purpose of the current study is to identify the predictors of user satisfaction with VPA from the perspective of technology acceptance. This study conducted partial least squares–structural equation modeling (PLS-SEM) to analyze 241 samples. The findings showed that satisfaction is significantly influenced by perceived ease of use, perceived usefulness, and perceived enjoyment. It was confirmed that the antecedents of technology acceptance have a significant effect on satisfaction with VPA. The results and this study will help researchers and practitioners to find new factors in explaining satisfaction.

Deep Learning Neural Network for Color and Volume Feature Extraction in Oral Cancer: Enhanced Co-occurrences

Faiza Shabbir, Abeer Alsadoon, Mohammad Zavid Parvez, Muhammad Siddiqi,
Thair Al-Dala'in, Fatima Furqaan, Razwan Mohmed Salah, Shayma Ismail Ali, Shahad Ahmed
School of Computing Mathematics and Engineering, Charles Sturt University, Australia;
School of Computer Data and Mathematical Sciences, Western Sydney University, Australia;
Kent Institute Australia, Sydney, Australia; Asia Pacific International College (APIC), Sydney, Australia;
The University of Duhok, Duhok, KRG, Iraq; Engineering Institute of Technology, Melbourne, Australia;
Peter Faber Business School, Australian Catholic University, Melbourne, Australia

Abstract: Local binary pattern (LBP) has been not effectively used for the texture classification of the oral tumour due to losing the dominant features. This study aims to enhance the classification accuracy of the early detection of oral tumour and improves the processing time by modifying normalized co-occurrences model by combining with the LBP. Moreover, the proposed model added the color and the volume to the extracted feature without damaging the other features. In the proposed study, deep leaning neural network was used for feature extraction and support vector machine (SVM) was used for classification using different sample images group of HSI images. The solution accuracy measured using probability score and processing time measured using total execution time where classification accuracy is 92% and approximately 50 to 60ms the processing time reduced.

Comparing Recurrent Neural Network Types in a Music Genre Classification Task: Gated Recurrent Unit Superiority Using The GTZAN Dataset

Eric Odle, Taukir Alam, Pei-Chun Lin, Amin Farjudian
Faculty of Information Engineering and Computer Science, Feng Chia University, Taiwan;
School of Computer Science, University of Nottingham, Ningbo P. R. China

Abstract: Due to the rapid growth in online media streaming services, accurate and efficient classification of musical genres has gained more prominence in the field of data science. In this paper, we evaluate common artificial neural network approaches—multi-layer perceptron (MLP), convolutional neural network (CNN), and three recurrent neural network (RNN) models, namely: Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), and Gated Recurrent Unit (GRU)— to maximize classification accuracy based on the GTZAN dataset. Our results show that the GRU model outperforms all other models tested with accuracy that is comparable to recent published work. Among the ten genres examined, our model can classify Hip-hop, Classical, and Metal music with the highest accuracy while exhibiting lower accuracy in distinguishing the music type of Country, Rock, and Pop.

Research on Intelligent Classification of Aggregate Based on Graph-based Image Segmentation

Weimin Qian, Caichao Zou, Jingen Yao, Renlei Liao, Xi Peng, Yanchao Wu, Haodong Zhan
Huzhou Xinkaiyuan Crushed Stones Co., Ltd, Huzhou, P. R. China;
School of Information Engineering, HuZhou University, Huzhou, P. R. China;
Zhejiang Province Key Laboratory of Smart Management & Application of Modern Agricultural Resources, Huzhou, P. R. China

Abstract: Aggregate gradation refers to the proportion of different particle sizes of aggregate, which is a crucial link in aggregate production. In this study, a total of 200 aggregate data sets and their labels were made based on aggregate samples provided by local gravel companies. Combined with the graph-based image segmentation algorithm, the aggregate was detected in real time, and the longest diameter of aggregate was obtained by making minimum circumscribed circle to the coloured area in the result graph, so as to judge the aggregate type. The experiment shows that the graph-based image segmentation algorithm can detect the aggregate accurately and efficiently, which is of great significance to the industry.

Optimism vs Skepticism in the Use of Intelligent Personal Assistant (IPA)

Hyeon Jo
Department of Strategic Planning, RealSecu, Busan, South Korea

Abstract: This research aims to explore the drivers of users' continuance intention in the context of intelligent personal assistants (IPA) from the perspective of optimism and skepticism. This study conducted a cross-sectional analysis by surveying Korean IPA users. Partial least squares–structural equation modeling (PLS-SEM) was used for analysis. The findings showed that system quality, information quality, and optimization are the main antecedents of continuance intention. The results indicated that skepticism does not have a significant effect on continuance intention. Based on the results, suggestions are made for researchers and service providers.

Human Activity Recognition Feature Extraction for Improved Classification Accuracy

Roberto Vazquez Ferrer, Jeffrey L. Duffany, Alfredo Cruz
Polytechnic University of Puerto Rico, Puerto Rico;
Ana G. Mendez University, Puerto Rico

Abstract: Machine learning became essential when increasing the availability of diverse types of sensor data. One of the most challenging is extracting features from the sensor data because the machine learning model suffers overlapping classification when using the raw sensor dataset. This paper presents a model that gives the electronic devices the notion of feature extraction of characteristics from the raw dataset from sensors like a wit-motion sensor for human activity recognition for disseminating information concerning intelligence activities to protect against such hostile actions. Feature extraction does strongly required before using machine learning for accuracy and Recall's incrementation of characteristics extracted from the raw dataset. Also, the following contributions provide a PCA assessment analysis that identifies overlapping classification based on the recall parameter and determines the feature to contribute more. The results show that the SVM recall parameter can help identify the classifier overlapping; the overlapping can observe using the PCA analysis. Finally, the feature extraction that scales the time series Magnitude of the acceleration by ten converts them into unique identify to detect unclassified activity. Moreover, the effect derives from creating a unique key that reduces the quantity of the original dataset representing new patterns.

Towards Adaptive Resilient Control of Evolving Environments

Nils Wenninghoff, Eric MSP Veith
OFFIS e.V. - R&D Division Energy, Oldenburg, Germany

Abstract: Deep Reinforcement Learning (DRL) applications are used in increasingly complex scenarios, such as Critical National Infrastructures (CNIs). The number of independent variables and external influences increases the challenge for DRL agents. The size of the sample space limits known DRL algorithms. CNIs additionally pose the challenge of transitioning to reality. The complexity forces many research studies to simplify simulations, resulting in a simulation-to-reality gap. In addition, simulation-based applications are often based on reward designs that cannot be implemented in reality, preventing online learning. This paper discusses two key problems and presents possible solutions including their realizations. Finally leading to a system that can operate in an online-learning scenario despite a high number of influencing variables.

An ISR Asset Planning Application with Intelligent Agents

Wilmuth Muller, Frank Reinert, Uwe Pfirrmann
Fraunhofer IOSB Institute of Optronics, System Technologies and Image Exploitation,
Karlsruhe, Germany

Abstract: This paper presents an approach for an optimal planning of ISR asset deploy-ment in order to satisfy the information needs of a commander. Based on the processes of information requirements management (IRM) and collection man-agement (CM), a two-step approach has been developed. In the first step, an op-erator assigns to each target on which reconnaissance or surveillance has to be performed a set of suitable assets. The operator may assign the suitable assets to a target either directly, based on his experience and knowledge, or supported by an interactive asset selection assistant component of the application, or supported by an intelligent multi-agent system, which generates automatically an asset assignment proposal. In the second step an optimal asset assignment and execution order is computed. The multi-agent system consists of three types of intelligent agents, a target agent representing the targets on which reconnaissance or surveillance has to be performed, the asset agents representing the assets available to the operator, and interface agents responsible for the communication with the other components of the application.

DRL-Driven Self Driving Network: Toward Optimized and Reinforced Forwarding Policies

Talha Ahmed Khan, Afaq Muhammad, Wang-Cheol Song
Department of Computer Engineering, Jeju National University, Jeju, Republic of Korea

Abstract: In recent years, networks have evolved towards remarkable softwarization, leading to a rapid increase in use of Software Defined Networking (SDN). The SDN is about separating the control plane from the data plane. It also helps Network Operators (NOs) to control the network from a central location and program northbound control applications to manage the data plane. However, all the controls are depended on setting priorities for traffic flows through the OpenFlow pipeline, not assuring e2e bandwidth or latency for a traffic flow. Nevertheless, OpenFlow enables the freedom to decide the policies using the control plane and it can be utilized to manage the modern network. It is the fact that future services like Massive IoT, Blockchain latency, and Metaverse demand sophisticated bandwidth-dependent policies. Modern services are sensitive to latency and cannot be satisfied using existing approaches to decide on data plane forwarding. Hence, this work proposes Deep Reinforcement Learning (DRL) to optimize the forwarding policies in the data plane. Due to the drastic amount of traffic and multiple types of services, it is highly required to decide on environment-driven reinforced policies rather than the existing algorithm fitting all services.

Exploring the Irish Healthcare: Identifying Needs for Services by Decision Tree Modelling

Anatoli Nachev
Business Information Systems, J.E. Cairnes School of Business & Economics,
National University of Ireland, Galway, Ireland

Abstract: This paper presents a case study that explores data form the Survey of Income and Living Condition (SILC), related to factors contributing to unmet healthcare needs in Ireland. Addressing a gap in previous research, we propose a machine learning modelling technique based on decision trees to identify and measure factors affecting the unmet healthcare needs in Ireland. Our experiments focus on three groups of factors: predisposing, enabling and needs. Measuring the model performance and variable importance, we can conclude that the top factors within each group are related to the overall health condition of the citizen combined with their abilities to afford paid healthcare services. This shows empirically, that the Irish healthcare system relates strongly unmet healthcare needs to the financial status of citizen and if they can afford private healthcare insurance.

Emotion Contagion by Geospatial Event Detection

Harshini Dutt Kuditipudi, Devinder Kaur
EEES Department, University of Toledo, USA

Abstract: Over the recent past, social media revolutionized the world due to its strong influence. The most common social media platforms as of now are Twitter, Facebook, Instagram, and Reddit. With over 300 million active user accounts, Twitter has one of the largest user bases. Opinions, emotions, trending topics, news are some of the contents of tweets. Twitter has one of the largest user bases; thus, making Twitter a major/popular social media platform for advertising and marketing, news updates, as well as social and data science. Emotional analysis and text mining techniques, such as Twitter API have been widely used to analyze or to mine tweets to extract useful information that would influence decision making. However, this paper uses a Text Sentiment analysis to achieve the purpose of this research, which is to examine emotion contagion by geospatial event detection.

Personalisation of d'Hondt's Algorithm and its Use in Recommender Ecosystems

Stepan Balcar, Ladislav Peska, Peter Vojtas
Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

Abstract: In the area of recommender systems, we are dealing with aggregations and potential of personalisation in ecosystems. Personalisation is based on separate aggregation models for each user. This approach reveals differences in user preferences, especially when they are in strict disagreement with global preferences. Hybrid models are based on combination of global and personalised model of weights for d'Hondt's voting algorithm. This paper shows that personalisation combined with hybridisation on case-by-case basis outperforms non-personalised d'Hondt's algorithm on datasets RetailRocket and SLANTour. By taking into account voices of minorities we achieved better click through rate.

ABC-based Optimal Hyper-parameter Tuning for Electric Load Forecasting

Yasaman Saadati, Leila Zahedi, Farid Ghareh Mohammadi, M. Hadi Amini
Knight Foundation School of Computing & Information Sciences, Florida International University, Miami, USA;
Sustainability, Optimization, & Learning for InterDependent Networks Laboratory (solid lab), Miami, USA;
School of Compuying, University of Georgia, Athens, Georgia, USA

Abstract: Electric load forecasting is crucial in the power systems to ensure generation-demand balance. The surge of machine learning, evolutionary algorithms and large-scale computing tools provided promising opportunities to build predictive models and enhance the prediction accuracy. The goal of deploying evolutionary algorithms is to optimize the prediction process in terms of forecasting performance by applying an efficient hyper-parameter tuning approach. In this paper, we exploit an automated hyper-parameter tuning method, a modified version of Artificial Bee Colony, OptABC, to solve a regression engineering problem. OptABC framework uses learning algorithms to tune the hyperparameters of two machine learning schemes (extreme Gradient Boosting and Random Forest) for enhancing the short-term load forecasting problem. The experimental results of this study verifies the effectiveness of the OptABC approach in finding the optimal hyper-parameter sets.

Proposal for a System to Automatically Create the Plot of a Novel and a Discussion of the Elements of Humanistic Writing

Misako Imono, Seiji Tsuchiya, Hirokazu Watabe
Department of Information Systems, Daido University, Nagoya, Japan;
Department of Intelligent Information Engineering & Science, Doshisha University, Kyoto, Japan

Abstract: In this study, we construct an automatic plot generation system for novels using a frame used in Hollywood (Beet-Sheet method) and investigate humanness based on the output of the system. The Beat-Sheet method is used as a frame to automatically construct the plot of a novel. The language used was English. The output plots are used to conduct a questionnaire for research and discussion. First, a tentative plot is generated and the names of people in the plot are listed. Various mechanisms are then applied, such as replacing one person's name with another or compressing the number of characters, to produce a final plot. This final plot is used to investigate humanness.

Application of Machine Learning to Radar System Identification

Jonathan Clement, Elham Khorasani Buxton
Department of Computer Science, University of Illinois at Springfield, Illinois, USA

Abstract: In this paper various models are explored for the purpose of providing multi-label classification based on sampled radar signal characteristics. A selection of radar systems are simulated to approximate realistic conditions that would be observed by a radar warning receiver. As there were no publicly available datasets appropriate to this task, a basic simulation was created and sampled to generate the required data. Several models were evaluated for radar identification including a simple multilayer-perceptron, a bi-directional recurrent neural network using LSTM, a convolutional neural network, a transformer model, MiniRocket, and XGBoost. These models were manually tuned to maximize validation accuracy. The convolutional neural network worked exceptionally well for this task and is a good candidate for further evaluation using a more robust environment which would include a larger set of radars with more details of their signal characteristics.

Efficient IoT Attack Detection in the Fog Layer with Machine Learning

Edna M. Chaar Santana, Jeffrey Duffany
Polytechnic University of Puerto Rico, Puerto Rico;
Ana G. Mendez University, Puerto Rico

Abstract: The use of Internet of Things (IoT) devices has grown significantly by 2021, making it more attractive for attackers to damage the system by stealing information, physically attacking the equipment, making it stop working correctly, blocking access, and many other possibilities of attacks. These attacks can occur on IoT equipment in businesses or an individual's residential environment. There is a crucial need to have an efficient intrusion detection system (IDSs) designed for this environment of IoT devices to mitigate attacks. This research will help find an efficient way to attack detection using Machine Learning methods on IoT devices in the fog layer. Three machine learning methods using classification algorithms will be studied, analyzed, and compared: K-nearest Neighbors, Random Forest, and Support Vector Machine. The most efficient way of attack detection will be identified.

Empirical Robustness Analysis of Learning to Incentivize Other Self-Interested Agents

Bengisu Guresti, Abdullah Vanlioglu, Nazim Kemal Ure
Istanbul Technical University, Sariyer Istanbul, Turkey

Abstract: Sequential Social Dilemmas are gaining attention in recent years. The current trends either focus on engineering incentive functions for modifying rewards to reach general welfare, or develop learning based approaches to modify the reward function by accounting for the impact of the incentive on policy updates. One of the most significant works in the learning based approach is LIO, which enables independent self- interested agents to incentivize each other by an additive incentive reward and demonstrates the method's success in several sequential social dilemma environments. We investigate LIO's performance under a variety of different setups in public goods game Cleanup in order to analyse its robustness against necessity of including inductive bias in incentive function, randomness in initial agent position with an option of asymmetric incentive potential, and assess its stability under frozen incentive functions after agents' explorations are reset. We observe and demon- strate empirically that LIO is indeed sensitive to these settings and it is not reliable for obtaining good incentives that would let the system stay stable when it is static. We conclude with some research directions that would improve the robustness of the method and incentive learning research.

A Collaborative Approach to Robocup Rescue Challenge

Abhijot Bedi, Shelley Zhang, Eugene Chabot
CIS, University of Massachusetts Dartmouth, North Dartmouth, Massachusetts, USA;
US Naval Undersea Warfare Center (NUWC), Division Newport, Rhode Island, USA

Abstract: Multi-robot coordination is required for a robot team to accomplish complex real-world tasks in a dynamic environment. RoboCup Rescue Simulation (RCRS) Testbed is a well-known environment to study robot teamwork. This paper describes a collaboration-based approach to the RCRS challenge. A domain-independent task modeling framework is adopted to model not only task decomposition but also collaboration requests. A Design-to-Criteria scheduler is used to schedule tasks accord-ing to a set of weighted criteria. This work would lead to bridge the collaboration re-search in multi-agent systems with the work in multi-robot systems where physical world constraints must be considered.

UAV Swarm Intelligence Simulation

Natalia S. Martinez Ojeda, Jeffrey L. Duffany
Polytechnic University of Puerto Rico, Puerto Rico;
Ana G. Mendez University, Puerto Rico

Abstract: Unmanned Aerial Vehicles (UAVs) are a type of aircraft that can be controlled remotely or programmed to fly in an autonomous way. They can be deployed individually or in squadrons depending on the situation. The security of the UAVs is critical to successful completion of their mission. Security countermeasures must be developed based real or perceived threats on UAVs from external attack. We assume a sophisticated adversary that has specific knowledge that can be used in his attacks. Previous work has reviewed a broad spectrum of plant and animal intelligence research to identify behaviors that might be applied to UAV security. A subset of those findings was chosen for more detailed analysis and simulation of the swarm behavior in attempt to develop deeper insights and the results are discussed herein.

Feature Engineering on Accelerometer Data for Human Activity Recognition Systems

Carla A. Varela-Rosa, Jeffrey L. Duffany
Polytechnic University of Puerto Rico, Puerto Rico;
Ana G. Mendez University, Puerto Rico

Abstract: Human Activity Recognition (HAR) Systems refer to procedures designed to recog-nize the actions of a human being based on a series of observations collected from sensor data. The challenge lies in preprocessing the raw sensor data via a method called feature engineering. This method tries to extract relevant pieces of information useful in the learning process of Machine Learning (ML) models. Another step constantly overlooked in HAR systems involves the methods for feature selection that attempt to tackle the phenomenon where models start to perform poorly with a great number of features. In this research, we investigated the impact of extracting key features from the time-domain and frequency-domain series of raw accelerometer data and their behavior through different configurations of segmentation and overlapping methods for the purpose of Human Activity Recognition. This was accomplished by selecting a public dataset composed of raw sensor data, extracting features, and exploring/validating a previously proposed brute-force approach for feature selection based on a K-Nearest Neighbors Classifier (KNC). Results show that for accelerometer based HAR systems, smaller window sizes are beneficial to the system's predictive performance, and that once key features are identified extra computations in overlapping percentages are not necessary.

Classification of NBA MVP Candidates Using Machine Learning

Matthew Schanilec, Robyn Smith, Isnino Harun, Malyun Ali, Matt Schoh, Mohammed Mahmoud
Department of Computer Science & Information Systems, Minnesota State University, Moorhead, Minnesota, USA

Abstract: N/A

Machine Learning-based Cyber Threat Anomaly Detection in Virtualized Application Processes

Jayesh Soni, Nagarajan Prabakar, Himanshu Upadhyay, Leonel Lagos
Knight School of Computing and Information Sciences, Florida International University, Miami, USA;
Applied Research Center, Florida International University, Miami, Florida, USA

Abstract: Intrusion-based detection systems spot traces of abnormal activities focused on the network and connected resources. Anomaly-based detection systems analyze events of applications for abnormal behavior based on the hypothesis that anomalies signify an indication of malicious events. Host-based systems frequently depend on various attributes of a process to describe the normal behavior of any process. Multiple malicious vectors can be launched on a process with different characteristics to infect it. We propose a two-step approach for host-based anomaly detection. First, we analyze ProcessList data structure and create Principal Component Analysis (PCA) features known as eigen traces used for training multiple one-class anomaly detection models. These multiple models allow different attributes of process data to be assessed from numerous and diverse standpoints. As the anomaly scores of these models vary significantly, combining the scores to a single value is often challenging. Therefore, we apply a majority voting approach for the final anomaly score as the second step. This final score measures the occurrence of a malicious event. In this study, we demonstrate the implementation of the proposed two-step approach using four different one-class classifiers: Mahalanobis Classifier, One-Class Support Vector Machine (OCSVM), Isolation Forest, and Dendogram based Agglomerative Clustering. We show that the proposed anomaly system improves the accuracy of anomaly detection.

Posture Estimation of Hugging-Type System by Motion Sensor Using Quaternion Neural Network

Chao Wang, Takehiko Ogawa
Mechanical and Electronic Systems Engineering Course, Takushoku University, Tokyo, Japan;
Department of Electronics and Computer Systems, Takushoku University, Tokyo, Japan

Abstract: Various types of neural networks extending to quaternions have been proposed. In particular, layered neural networks using neurons that realize quaternion geometric operations have been applied to transformations and color image processing in three-dimensional spaces. Estimating the three-dimensional posture of an object is important for analyzing human motion and computer graphics. Alternative, we have proposed a hugging-type system for medical support. This system measures the state of the human while hugging the sphere where the sensor is attached. This study considers using the quaternion neural network to estimate the posture of an object. We propose quaternion neural network to estimate the posture of the hugging-type system.

CALIME: Causality-Aware Local Interpretable Model-Agnostic Explanations

Martina Cinquini, Riccardo Guidotti
University of Pisa, Pisa, Italy

Abstract: A significant drawback of eXplainable Artificial Intelligence (XAI) approaches is the assumption of feature independence. This paper focuses on integrating causal knowledge in XAI methods to increase trust and help users assess explanations' quality. We propose a novel extension to a widely used local and model-agnostic explainer that explicitly encodes causal relationships in the data generated around the input instance to explain. Extensive experiments show that our method achieves superior performance comparing the initial one for both the fidelity in mimicking the black-box and the stability of the explanations.

Semantic Enrichment of XAI Explanations for Healthcare

Luca Corbucci, Anna Monreale, Cecilia Panigutti, Michela Natilli, Simona Smiraglio, Dino Pedreschi
Department of Computer Science, University of Pisa, Pisa, Italy;
ISTI-CNR Pisa, Italy

Abstract: Explaining black-box model decisions is crucial to increase doctors' trust in AI-based clinical decision support systems. However, current eXplainable Artificial Intelligence techniques usually provide explanations that are not easily understandable by experts outside of AI. Enriching the explanations with relevant clinical information concerning the health status of a patient would increase the ability of human experts to assess the reliability of the AI decision. Therefore, in this paper, we present a methodology that aims to enable clinical reasoning by semantically enriching AI explanations. Starting from a medical AI explanation based only on the input features provided to the algorithm, our methodology leverages medical ontologies and NLP embedding techniques to link relevant information present in the patient's clinical notes to the original explanation. Our experiments, involving a human expert, highlight promising performance in correctly identifying relevant information about the diseases of the patients.

brainIT: A Generic IT Core Mechanism for Continuous Growth-Flow in Dynamic Chaotic Context

Muhammed Akif Agca
CE Department TOBB University of Economics & Technology, TOBB ETU, Ankara,
Turkiye Luxembourg Institute of Science & Technology / LIST, Esch-Sur-Alzette, Luxembourg

Abstract: Brain CAPEX (Capital Expenses) is for free to the human being but the OPEX (Operational Expenses) is not. Since, the fluctuations on critical nutrition for brain makes it complicated to grow via optimal path as continues progress due to the chaotic OPEX changes. Fortunately, intelligent systems are able to adapt the change dynamically up to varying chaotic context, by keeping trustworthiness of the whole system via available distributed resources and algorithms. However, increasing number of nodes in the system inflates complexity of swarm behavior due to computation and memory limitations. Drastic progress saved in the emerging edge devices, can enable to produce innovative trusted AI/ML algorithms at run-time, which can help to make massive analytics at the edge nodes in (near) real time. In spite of this, keeping the system resilient require real-time updates in different system layers. As another critical milestone, increased scalability and faster in memory processing speed can be accomplished via big data technologies and ledger base chained structures in some manner. In order to keep high performance of the total system, mission/safety/operation critical applications require to be verified by critical check-points. Thereby, end-to-end trust mechanism and swarm controller methods can improve trusted scalability of the intelligent systems analytical functions and resources. So that, the dynamic holistic views can ensure trustworthiness in chaotic context with the brainIT generic IT core mechanism for continuous growth in massive-chaos, which ensures to keep local/global legal constraints-based risk minimization via 5G connected hybrid-cloud systems within the observed socio-dynamic parameters with minimized optimal OPEX costs.

Quantifying the Increase of Vegetation Area along the Brink of the Taklamakan Desert with Deep Learning Model and Satellite Image

Zhiyi Li, Dali Wang, Daniel Riccuito
Department of Computer Science, Virginia Tech, Blacksburg, Virginia, USA;
Environmental Sciences Division, Oak Ridge National Lab, Oak Ridge, Tennessee, USA

Abstract: Large-scale man-made revegetation has been reported to reduce desertation in northwest China. We presented a method to quantify the vegetated area across large regions by tracking the phenological cycle of vegetation with satellite images (Normalized Difference Vegetation Index (NDVI) dataset from NASA Moderate Resolution Imaging Spectrometrometer (MODIS)) and a recurrent neural network. This study showed that the vegetated area has increased by 6395.5 square km from 2012 to 2021 while the NDVI distribution remains unchanged. Based on the location of these increased vegetated areas and local reports, the steady increase of the vegetated area is believed as the direct outcome of the man-made re-vegetation effort in the past decades.

Evolutionary Algorithms for Automated Federated Learning

M. Hadi Amini, Ahmed Imteaj, Yasaman Saadati, Farid Ghareh Mohammadi
Knight Foundation School of Computing & Information Sciences, Florida International University, Miami, USA;
Sustainability, Optimization, & Learning for InterDependent networks laboratory (solid lab),
Florida International University, Miami, Florida, USA;
School of Computing, University of Georgia, USA

Abstract: The extensive growth of data generation, sensing, and collection at the edge devices lead to inconvenience in terms of data storage, management, and ensuring privacy protection of associated data while sharing with central servers. This motivated development of a distributed machine learning (ML) paradigm, Federated Learning (FL) that leverages training a model without exposing any data utilizing the local computational and data resources of the edge devices. However, the existing FL research mostly focused on optimizing model quality, communication overhead, and privacy, whereas, limited or less efforts were put into analyzing the significance of data features and tuning of local models' hyperparameters. In this paper, we propose to develop a holistic framework that uses evolutionary computing for optimal hyperparameter tuning and feature selection optimization of our novel FL algorithms, i.e., we are planning to provide researchers with a comprehensive roadmap that can be used to enable automated FL. Given the fact that our algorithms are widely used for IoT, robotics, critical infrastructure resilience, and healthcare, an automated platform will add another layer of practicality and enable real-world integration of these algorithms.

Generalized Churn Classification Across Multiple Business Domains

Maryam Naveed, Arslan Shaukat, Usman Akram, Sajid Gul Khawaja
National University of Sciences and Technology (NUST), Islamabad, Pakistan

Abstract: For any organization, customers are the basis for company success. Research in Customer Relationship Management (CRM) shows that it is more beneficial to retain customers. It guarantees a higher return than acquiring new ones at five times the cost. For this purpose, organizations target minimal churning. Churning is defined as any customer ending a subscription or stopping the use of a service provided by an organization. This research aims to provide a generalized system that includes pre-processing and feature selection, which can be utilized with different components and business rules to identify customers on the verge of churning. A centralized hybrid algorithm has been devised to identify possible at-risk customers. The gap created when a researcher has to rely on a hit and trial method to locate the best possible algorithm to solve their problem has been addressed. IBM Watson and Cell2Cell and locally sourced data have been used with classifiers such as Support Vector Machines, GP-AdaBoost, and Random Forest. The pipeline includes SMOTE-ENN sampling, RFE feature selection, and normalization techniques. We have achieved the highest accuracy of 0.984 on IBM Watson and 0.994 on the Cell2Cell dataset with this setup. The locally sourced dataset has not been used in previous research. Hence, it has been used as scoring data on which we have achieved an accuracy greater than 0.99. The results achieved on the two benchmark datasets using our proposed system are competitive compared to prior literature.

Research on the Setting of Text-based Policy Agenda for the Improvement of Human Rights and Treatment of Social Workers

Seungjae Lee, Soongoo Hong, Sungho Park, Heiwon Lee, Seunghyeon Lee, Kangbae Lee
Data Analysis & Artificial Intelligence Research Lab., Dong-A University, Seo-gu, Busan, Republic of Korea;
Smart governance Research Laboratory, Dong-A University, Gudeok-ro, Seo-gu, Busan, Republic of Korea

Abstract: The role of social welfare workers is emphasized in Korea's modern society, but human rights and treatment are not good. So many people suffer from burnout and change jobs. Until now, ordinances for social welfare workers have been conducted, but the level of implementation is low and the feeling of improvement is also low at 4.23 out of 10. To solve this problem, it is necessary to understand the hardships of social welfare workers. In this research, a method is proposed for understanding the hardships of social welfare workers and setting effectively a policy agenda using online text data and text analysis techniques that have not been conducted before. Frequency analysis, topic analysis, and similar word analysis are conducted, and policy agendas are set by conducting mutual comparison and analysis to compensate for the limitations of each analysis method. This research is meaningful in that the policy agenda was set using the text analysis method to improve the human rights and treatment of social worker.

Non-Atomicity Detection of Natural Language Requirements Based on Transfer Learning

Shaobin Li, Yihan Cai, Qinggu Li, Bo Wang, Xinggang Yang, Weidong Yang
Data Science Research Institute, City University of Macau, Macau, P. R. China;
School of Computer Science, Fudan University, Shanghai, P. R. China;
Big Data Centers Gree Electric Appliances, Inc. of Zhuhai, Zhuhai, P. R. China

Abstract: Most system requirements are written in natural language, and the quality of natural language requirements greatly affects subsequent development. Focusing on the problem of requirement quality, we comprehensively took the definitions of requirement defects in related work into consideration, and collected public-available real-world requirement data, on which the experiments were carried out through the existing rule-based model. After finding that existing rule-based defect detec-tion models are not effective in dealing with non-atomicity in real-world data, we proposed a defect detection method based on BERT (Bidirectional Encoder Representations from Transformers) language model and convolutional neural network. Further experiments about model fine-tuning, preprocessing, and comparison on smaller-scale dataset were also conducted. The results proved that our model could effectively improve the effect of detecting non-atomicity, and it also outperformed traditional machine learning models.

Cost-Effective Resource Provisioning of Cloud Computing via Supervised Machine Learning

Daniel Andresen, Mohammed Tanash, Caughlin Bohn, William Hsu
Computer Science Department, Kansas State University, Manhattan, Kansas, USA;
Holland Computing Center, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

Abstract: In this paper, we investigate the impact on cost effectiveness of employing Machine Learning techniques for predicting job resources (memory and time) in terms of resources and cost provisioning for running HPC workloads on the cloud. We evaluated our AMPRO-HPCC tool by comparing the run time and cost of 4.46 million jobs covered 2018-2021 years derived from Kansas State University HPC cluster (BEOCAT) logs. We found that our Machine Learning tool reduced the average cost of running jobs on the cloud by up to 39% and decrease the average running time by up to 39%.

Automated Acne Assessment

Yihang Li, Alex Kot
School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

Abstract: Deep learning methods are gaining popularity in medical image analysis. Development of deep learning based on facial analysis methods suffers from lack of data and existing approaches are sensitive to pose, light, ethnicity, and occlusion, which causes large domain gap. To solve this problem, we proposed a central difference double channel network. In this model, central difference convolutional neural network is deployed to capture gradient information of acne towards common skin. Label distribution learning and multi-task learning skills are also implemented to enhance the learning ability. We propose an adapted meta learning network as well as transfer learning skills to improve the performance of the adapted meta learning network on a selfie dataset.

Convolutional Neural Networks Based Algorithm Using Compressed Sensing for Butter Price Forecasting

Jaroslaw Malczewski
WSB University Poznan, Poland

Abstract: The latest studies have compellingly argued that Neural Networks (NN) classification and prediction are in the right direction for forecasting. It has been proven that NN are suitable models for any continuous function. Moreover, these methods are superior to conventional methods, such as: Box-Jenkins, AR, MA, ARMA, or ARIMA. The latter assume a linear relationship between inputs and outputs. This assumption is not valid for skimmed milk powder (SMP) forecasting, because of nonlinearities, which are supposed to be approximated. The traditional prediction methods need complete date. When values are missing, mispredictions may result. Moreover, the traditional prediction methods operate using linear dependencies. As such, in many non-artificial-intelligence-based prediction procedures, the main assumptions are linearly based, and not complex joint distribution based. The non-AI-based techniques regularly handle univariate-like data only. This assumption is not sufficient, because many external factors may possibly influence the time series. Unfortunately, due to the complexity of SMP price series, typical stochastic-based algorithms are unable to deliver satisfactory results. In this paper, a novel, hybrid, artificial-intelligence-based forecasting model is presented. It should be noted that any Artificial Neural Network (ANN) approach can be strongly affected by the relevancy and "clarity" of its input training data. To enhance forecasting results, initial training data is needed. In this paper, a modified training paradigm based on the compressed sensing (CS) theorem is proposed. In the proposed methodology, the CS framework assumes price series data to be sparse and noisy. The presented procedure utilizes CS methodology, which assumes noisy trends are incomplete signals for them to be reconstructed using CS reconstruction algorithms. Denoised trends are more relevant in terms of NN-based forecasting models' prediction performance. Empirical results have revealed robustness of the proposed technique. These results show how superior the technique is.

Weighted Ensembles of Multiple Class Rankers

Jonathan I. Lee, Sung-Hyuk Cha
Computer Science Department, Pace University, New York, USA

Abstract: This paper proposes a novel approach to ensemble pattern recognition, by combining weighted-vote aggregation with rank ordering, and finds it superior to each approach in isolation. Such ensemble methods are of great interest in machine learning and pattern recognition. Generally, research in ensemble methods had moved in two different directions: weighted aggregation on one hand, and ranking on the other. Here, we propose to employ both together, a method using weighted, preferential voting with multiple rankers. This paper includes a brief experiment, or proof of concept, using convolutional neural network (CNN) models with differing convolutional bases on a public image data set and demonstrates the higher performance of the proposed hybrid method.

Approach to the Narrative Transformation Through Cultural Love and Sex Knowledge

Takashi Ogata
Faculty of Software and Information Science, Iwate Prefectural University, Japan

Abstract: The theme of love and sex is one of the most important topics in narratives and literature. The author has focused on love & sex in the context of narra-tive generation system study based on artificial intelligence, cognitive sci-ence, and robot technologies. This paper shows the basic concept and macro architecture of the system that transforms an input narrative into another narrative, including human relationships, acts and thoughts, and representa-tion regarding love & sex. Moreover, the system uses the knowledge of Jap-anese narratives and literature from the viewpoint that love & sex knowledge includes cultural characteristics.

Hybrid Convolutional Neural Networks and Modified GARCH Model Based the European Union's Emissions Trading Scheme Prices Forecasting

Krzysztof Malczewski, Zbigniew Krysiak, Adrian Markiewicz
Institute of Information Technology, Warsaw University of Life Sciences, Warsaw, Poland;
Department of Corporate Finance Management, Warsaw School of Economics, Warsaw, Poland;
Faculty of Management and Security Studies, University of Social Sciences, Warsaw, Poland

Abstract: Carbon prices follow a stochastic process of complex time series with nonstationary and nonlinear characteristics, making forecasting CO-2 emission prices an important and difficult task for policymakers and market participants. Existing literature has focused on highly precise point forecasting, but in most cases, it cannot correctly solve the uncertainties associated with carbon price datasets. Because the volatility of the European Union's Emissions Trading Scheme (EU ETS) has time-series characteristics, as well as long-term memory, volatility aggregation, asymmetry, and nonlinearity, this study proposes an ETS volatility prediction model by combining convolutional neural networks (CNN), and long short-term memory (LSTM) network and generalized autoregressive conditional heteroscedasticity (GARCH) family models. The proposed Convolutional Neural Networks-based methodology assumes sparse and noisy input data. The presented method employs Compressed Sensing methodology, which assumes that noisy time series are incomplete signals that can be reconstructed using CS reconstruction algorithms. Denoised training sets are more relevant in terms of the prediction performance of NN-based forecasting models. The proposed technique is robust, according to empirical results.

Considering the Story Generation Method Using Coloring Techniques Included in a Story

Jumpei Ono, Takashi Ogata
Faculty of Software and Information Technology, Aomori University, Tokyo, Japan;
Faculty of Software and Information Science, Iwate Prefectural University, Iwate, Japan

Abstract: We propose a story-generation system with story techniques included in the sto-ry. In this study, a story is a structure that represents the content of a narrative. The story technique is a procedure that extends or edits the story's structure. Al-ternatively, the story technique can be used as a function. We propose a system based on the concept of an automatic story-generation game. The proposed system uses a mechanism for generating an entire story. The characters in the gener-ated story use story techniques. Based on changes in the story, the characters extend or edit the story using their own techniques. We use "coloring" as the story technique. The proposed mechanism attempts to generate a story having the ability to modify the story that is not expected by the user.

Enhanced Dengue Outbreak Prediction in Tamilnadu using Meteorological and Entomological Data

Varalakshmi M, Daphne Lopez
VIT Vellore, TamilNadu, India

Abstract: This paper focuses on studying the impact of climate data and vector larval indices on dengue outbreak. After a comparative study of the various LSTM models, Bidirectional Stacked LSTM network is selected to analyze the time series climate data and health data collected for the state of Tamil Nadu (India), for the period 2014 to 2020. Prediction accuracy of the model is significantly improved by including the mosquito larval index, an indication of VBD control measure

SESSION/WORKSHOP
XXII Technical Session on Applications of Advanced AI Techniques to Information Management for Solving Company-Related Problems

Co-Chairs: Dr. David de la Fuente and Dr. Jose A. Olivas***
**University of Oviedo, Spain*
***University of Castilla - La Mancha, Spain*

Solving Rescheduling Problems in Dynamic Permutation Flow-Shop Environments with Genetic Algorithm

P. Valledor, A. Gomez, J. Puente, J. Parreno
Escuela Politecnica de Ingenieria Gijon, Campus de Viesques s/n, Gijon (Asturias), Spain

Abstract: The aim of this paper is to analyse, model and solve the rescheduling problem in dynamic permutation flow shop environments while considering several criteria to optimize. In this paper, a restarted iterated Pareto greedy (RIPG) metaheuristic is used to find the optimal Pare-to front. To demonstrate the appropriateness of this approach, the algorithm is applied to a benchmark specifically designed in this study, considering three objective functions (makespan, total weighted tardiness and steadiness), and three classes of disruptions (appearance of new jobs, machine faults and changes in operational times).

Industry 4.0 as a Driver of Financial and Non-financial Performance

OAL. Garcia, I. Fernandez, R. Pino, B. Ponte
Escuela Politecnica de Ingenieria Gijon, Campus de Viesques s/n, Gijon (Asturias), Spain

Abstract: Industry 4.0 offers various advantages to industrial development and organizations are implementing it as a strategy to improve their business performance. Implementing industry 4.0 can be costly and complicated due to the limited re-sources most organizations have, as it integrates various technologies, changes at the structural and organizational level that can generate additional expenses. This makes organizations want to see the benefits of their adoption before purchasing such technologies. Through an empirical analysis, this study evaluates the use of technologies related to industry 4.0 with measures of financial and non-financial performance. With a multivariate analysis of canonical correlation, the study found that the most outstanding benefits in terms of financial performance (profit margin and sales growth) and non-financial performance (customer relation-ship, innovation of their own products and services, assurance of product quality) are mostly driven using technologies related to cybersecurity, Cloud computing, RFID and BigData.

Blockchain's Distributed Ledger on Supply Chain: Discourse on the Method

S. Fernandez-Vazquez, R. Rosillo, N. Alvarez, D. De la Fuente, M. Luna
Escuela Politecnica de Ingenieria Gijon, Campus de Viesques s/n, Gijon (Asturias), Spain

Abstract: Blockchain enables businesses to comprehend their supply chain and interact with customers using authentic, verifiable, and unchangeable data. As a result, there is a growing interest in how this new technology may affect the supply chain, particularly in terms of measuring and valuing it. From a supply chain perspective, we investigate the growing interest in blockchain technology in this paper with the goal of identifying key players and classifying the conceptual framework of existing research. To do this, we combine conventional bibliometric analysis in VOSviewer with text mining analysis based on Natural Processing Language, using evaluative and relational methodologies (NLP).

An Ontology and Taxonomy Based Model for Dealing with Radicalization in Social Media

Andres Montoro, Jose A. Olivas, Francisco P. Romero, Jesus Serrano-Guerrero,
Javier Morejon-de-Giron, Enrique Larriba-Gonzalez, Antonio Ramos
Computer Science Department, University of Castilla-La Mancha, Paseo de la Universidad, Ciudad Real, Spain

Abstract: Among its many uses, social media, and more specifically social networks, have served as a space for spreading propaganda of radical ideals and promoting terrorism. The use of social media for terrorist purposes creates opportunities and challenges in the fight against terrorism. As part of the Phoenix platform being developed by Mollitiam Industries for Homeland Security (HLS) and Homeland Defense (HLD), we have developed a prototype for the detection of potential radicals on social media. The paper is divided into two main parts, first the characterization of the radical profile using a taxonomy and then on the semantic categorization of a particular radical profile from the development of a domain ontology. To reason about taxonomy we use fuzzy inference. The proposal is implemented and tested in the Phoenix framework with the aim of scaling it to different domains.

SESSION/WORKSHOP
Innovation Technology and Theory for Metaverse Transformation

Chairs: Prof. Young-Ae Jung and Prof. Charlie (Seungmin) Rho***
**Division of Information Technology Education, Sunmoon University, Asan, Republic of Korea*
***Department Software, Sejong University, Gwangjin-gu, Seoul, Republic of Korea*

Transmission Model for Blockchain-Based Regional Energy Internet

Junho Kim, Jaeeun Kim, Mucheol Kim
School of Computer Science and Engineering, Chung-Ang University, Seoul, Korea

Abstract: With the development of decentralized renewable energy, many prosumers have emerged, and the Regional Energy Internet (REI) has emerged. As the number of entities participating in energy transactions increases and trans- action types become more diverse, it is necessary to minimize the problems of information inconsistency, difficult-to-build trust system, and power deviation waste and cost increase caused by the centralized smart grid model. Therefore, in this paper, we propose a blockchain-based energy information transmission model with improved energy efficiency and security. The model consists of two core layers: blockchain and agent. Blockchain designs smart contracts to support data integrity and user verification in energy networks. The agent performs data integrity verification between the blockchain and the grid and supports transac- tions with enhanced security between the prosumer and the grid.

Deep Learning-based Power Measurement Model for Effective Energy Consumption

Hyunwook Yu, Mincheol Shin, Mucheol Kim
School of Computer Science and Engineering, Chung-Ang University, Seoul, Korea

Abstract: Complex operations across a variety of industries are driving demand for HPC. Since the HPC environment uses a lot of energy, there are much re-search on energy management policies. Efficient energy management policies should be based on accurate power consumption measurements. However, most research show limits in accuracy by calculating and multiplying weights using hardware performance and I/O counters. In this paper, we propose a model that can accurately measure the power consumption with deep learning. We collect the power of the CPU and the performance counter and select features by analyzing the correlation between counter and power. We propose a deep learning-based power measurement model trained with the features.

Research on the Structure of a Development Environment Corresponding to the Color Reproduction Pipeline of the Headset Display for Virtual Reality

Hyun Suh Kim, Jung Yoon Kim
School of Photography and Video, Kyung-il Univesity, Gyeongsan, Korea;
Graduate School of Game, Gachon University, Seongnam, Korea

Abstract: N/A

Adaptive Data Management based on Data Usage Pattern for Large-scale Distributed Storage System on WAN

Jaehwan Lee, Kyoungchan Kim, Sangoh Park
School of Computer Science and Engineering, Chung-Ang University, Dongjak-gu, Seoul, Korea

Abstract: Recently, research on expanding storage to meet global data demands is being actively conducted. In order to overcome the limitation of the storage capacity of data centers, it is necessary to establish distributed storage through connection between data centers. Since WAN connections between data centers exhibit high latency, the location where data is stored is important in order to reduce the impact of latency. However, the existing distributed storage system technology does not take into account the transmission latency between data centers. In this paper, we propose a distributed storage technology that provides low-latency data access even in a WAN environment by storing data in a location with a high access frequency based on user's data access information.

A Survey on Fault Tolerance Architectures in Cloud Systems

Mingyu Jo, Sangoh Park
School of Computer Science and Engineering, Chung-Ang University, Dongjak-gu, Seoul, Korea

Abstract: N/A

A Video-based Person Re-identification Using an Optimized-Deep Learning Model and Pseudo Labeling

Eliza Batool, Muazzam Maqsood, Maryam Bukhari, Seungmin Rho
Department of Computer Science, COMSATS University Islamabad, Attock Campus, Pakistan;
Department of Industrial Security, Chung-Ang Univrsity, Seoul, South Korea

Abstract: Person Re-identification (Re-ID) refers to recognizing a person as the same person even if they appear in a different camera. It is becoming increasingly important in computer vision for real-world surveillance applications. Deep learning methods are currently widely used in person-identification challenges. However, for proper training, these models need a massive amount of data. To tackle the difficulty of insufficient data, previous studies used pseudo labelling methods with deep learning. However, one another problem is the deep learning model's hyperparameters. As a result, in this study, we tune the parameters of deep learning models to make it optimal along with pseudo labeling technique to boost performance even more. More precisely, data for training the model is gathered from two major datasets: DukeMTMC-VideoReID and MARS. Deep learning model namely ETAP-Net generates an initial model that serves as the foundation for the development of pseudo labels for unlabeled data. Later on, the model is optimized with best set of parameters to improve the performance. The suggested method is evaluated using the Cumulative Matching Characteristics (CMC) and mean Average Precision (mAP) criteria to truly determine how the model with best set of parameters assists in achieving correct labels for unlabeled data.

Crawling Analysis of Industrial Technology Leakage in the Darknet

Youngjae Kong, Hangbae Chang
Department of Security Convergence, Graduate School, Chung-Ang University, South Korea;
Department of Industrial Security, Chung-Ang Univrsity, Seoul, South Korea

Abstract: N/A

Efficient Security-level Text Classification Method with Small Imbalanced Dataset Using RNN Model

Yuna Han, Hangbae Chang
Department of Convergence Security, College of Business & Economics, Chung-Ang University, Seoul, South Korea;
Department of Industrial Security, Chung-Ang Univrsity, Seoul, South Korea

Abstract: Classifying the security-level document has been an area in the industry to prevent confidential data leakage. Previous studies on automated-based classification methods, such as Naive Bayes and support vector machine (SVM), have enhanced accuracy and time efficiency. However, these methods can be challenging in the real world when using small and unbalanced datasets. To solve this problem, this paper proposes a novel security-level text classification based on a recurrent neural network (RNN). We first collect the small and imbalanced real-world WikiLeaks dataset and then perform data preprocessing such as removing stop-words, reducing the dimension of feature vectors, and so on. We develop the optimized RNN model via the sequence of text by applying a dropout layer and L2 regularization. Experimental results show that our proposed scheme can achieve higher accuracy than other controlling methods, as well as existing machine learning techniques and deep learning models. Through these experiments, we conclude that our proposed scheme could be more effective (99.60%) in converging than other tuning ways and hence outperform the keyword-based mechanism.

A Case Study on the Industrial Security Expert Education

Harang Yu, Hangbae Chang
Department of Convergence Security, College of Business and Economics, Chung-Ang University, Seoul, South Korea;
Department of Industrial Security, Chung-Ang Univrsity, Seoul, South Korea

Abstract: As technology develops and environment changes, new security threats emerge in diverse ways. In order to cope with rapidly changing security threats and issues, it is necessary to establish appropriate curriculum and educate future security experts. However, current security education is comparatively focused on technical view of security rather than having a whole perspective toward security. With timeliness characteristics, security not only needs academic knowledge but also field-based experience and requirements in order to cope with continuously changing environment. In this study, manufacturing industry is considered as the most affected industry from new technology and designed and conducted a security curriculum for future security R&D experts. It was intended to not only learn an academic knowledge regarding security, but also acquire field-based experience, know-how, etc. in order to satisfy as a security experts' requirements. As a result, students who participated in corresponding curriculum resulted in getting a security related job, conduct an internship, etc. which proved to show the importance of academic knowledge and field-based curriculum.

Research on Security Management Strategies of Core Technology Leakage due to Brain Drain

Yurim Choi, Hangbae Chang
Department of Convergence Security, College of Business & Economics, Chung-Ang University, Seoul, South Korea;
Department of Industrial Security, Chung-Ang Univrsity, Seoul, South Korea

Abstract: With the advent of 'Pax Technica' in which possessing technology leads to corporate competitiveness, technological leakage is carried out through hacking and M&A, as well as leakage of core personnel who develop and manage core technologies (brain drain). Considering these technology leakage methods, we propose strategies to prevent technology leakage due to brain drain. Thus, we design strategies for each working step to prevent brain drain through analyzing related previous studies and literature on brain drain and security countermeasures.

A Study on Estimation of Spatial Structure Using Hole Restoration of Indoor Floor Point Cloud

Jaeseok Yoo, Jeongha Lee, Sanghyun Seo
Graduate School of Advanced Imaging Sciences, Multimedia & Film, Chung-Ang University, Seoul, South Korea;
School of Computer Art, Chung-Ang University, Anseong, South Korea;
College of Art & Technology, Chung-Ang University, Anseong, South Korea

Abstract: In this paper, we introduce a method of estimating the spatial structure by restoring a hole in the floor surface in a virtual indoor space point cloud constructed based on real-world visual information. The floor surface is a very important clue in modeling real-world building structures. In particular, the corner information on the floor surface plays a role in connecting the structure of the building, and through this, the overall structure of the space can be estimated. However, in the case of an indoor space, there are objects (chairs, desks, beds, etc.) on the floor. This occludes the boundary of the floor surface, and the point cloud of that area is lost. This is a problem to be solved in modeling a 3D virtual space, and it is a very difficult task. We restore information on the floor surface through GAN. As a result, it is possible to know the exact corner information of the floor surface, and through this, the structure of the space can be estimated and modeled.

Irregular Quadrilateral Region Tracking using ARIMA for Dynamic Projection Mapping

Bumsoo Kim, Changhun Jung, Yonghoon Jung, Sanghyun Seo
School of Computer Art, Chung-Ang University, Anseong, South Korea;
Graduate School of Advanced Imaging Sciences, Multimedia & Film, Chung-Ang University, Seoul, South Korea;
College of Art & Technology, Chung-Ang University, Anseong, South Korea

Abstract: In this paper, we introduce a dynamic projection mapping system that can predict the next trajectory and region of rectangular object based on ARIMA, a time-series model. It is designed to overcome the system latency by tracking with general device. This can offset the problem of inaccurate matching between the projected region and the real-world interest region of interest caused by system latency. In experiments, our system performed DPM on rectangular object with irregular quadrilateral region in the captured image. For the evaluation, we utilize the irregular IoU tool and compared the predictive models corresponding to ARIMA order.

Retinal Optical Coherence Tomography Image Classification Based on Hybrid Ensemble Deep Learning

Kuntha Pin, Jung Woo Han, Yunyoung Nam
Department of ICT Convergence, Soonchunhyang University, Asan, South Korea;
Department of Ophthalmology, Bucheon Hospital, Soonchunhyang University College of Medicine, Bucheon, South Korea;
Department of Computer Science and Engineering, Soonchunhyang University, Asan, South Korea

Abstract: The OCT image has been known as an essential modality in clinical and used by ophthalmologists to diagnose various retinal diseases. Manual analysis is time-consuming, and it might be hard to distinguish various features of diseases even with more experience. Therefore, this study proposed an automatic OCT image classification to diagnose retinal disease through hybrid deep learning and ensemble machine learning. Transfer learning with convolutional neural network (CNNs) model is used to extract a thousand features from OCT images. Then the extracted features are used as features for training machine learning methods. By integrating two different tasks, the global features extracted by deep learning are classified by voting classifiers to boost the performance of five-class retinal image classification. The ensemble classifier is a fuse of two machine learning classifiers, K-Nearest Neighbors (KNN) and XGBoost. Furthermore, to form the proposed method more applicable and accusable, the deployment of the retinal OCT image classification on a web server achieved the response time in seconds. In addition, we also consider various CNN models to achieve better performance, and the result of the proposed method obtained a better performance of 97.68% in the InceptionResnetV2 model.

The 18th International Conference on Data Science
(ICDATA'22: July 25-28, 2022, USA)
https://icdata.org
https://american-cse.org/csce2022/

Random Graphs in DGL

Peter M. Maurer
Department of Computer Science, Baylor University, Waco, Texas, USA

Abstract: Our recent upgrade to the DGL system for generating random data includes sev-eral features to generate random graphs. These features permit random graphs to be inserted into user-defined C++ variables, and then to be used in any manner desired. There is a strong coupling between the DGL features and the software to be tested, allowing both to work in tandem. We are able to generate graphs ac-cording to several popular models. We are able to generate purely random graphs, connected graphs, bipartite graphs, directed graphs, directed acyclic graphs and trees. We are also able to generate binary trees, search trees and red/black trees. We outline the algorithms used to generate each type of graph.

Assessment of Trust Models in Supply Chain Traceability Data Management

Oratile Leteane, Yirsaw Ayalew
Department of Computer Science, University of Botswana, Botswana

Abstract: Several trust models have been developed for enhancing trustworthiness in different supply chain management systems. However, most of the models have not been used when developing supply chain traceability systems. This study investigates the relevance of these existing trust models in addressing trust issues in traceability data. This is done by analyzing existing supply chain computational trust models presented in the literature in the past ten years, from the year 2012 to 2022. The trust models were analyzed based on the trust metrics used, how the metrics were identified and whether such metrics measure the trustworthiness of supply chain traceability data. Analysis shows that there are a variety of metrics suggested by different researchers. While these metrics vary, most of the models use supply chain transaction success and failure rates to determine trust values. However, the results show that trust metrics were not developed to address trust issues in traceability data. This study further suggests a systematic approach to determining trust metrics in supply chain management.

A Recommendation System for Short Recipe Videos with Supplementary Cooking Operation

Takuya Yonezawa, Yuanyuan Wang, Kazutoshi Sumiya, Yukiko Kawai
Kyoto Sangyo University, Kyoto, Japan;
Yamaguchi University, Yamaguchi, Japan;
Kwansai Gakuin University, Hyogo, Japan

Abstract: We propose an interactive cooking support system for short recipe videos by extracting and weighting cooking operations for each cooking genre based on user browsing behavior. The system then recommends various supplementary recipe videos based on the weights of cooking operations and user browsing behavior. Also, the supplementary recipe videos can be dynamically changed based on the user browsing behavior. As a result, users can intuitively and easily understand cooking operations suited to their cooking favorites. In this paper, we verified the effectiveness of the weighting of cooking operations by 60 subjects.

An Integrated Dataset of Public Natural Language Requirements with Labels of Defects

Yihan Cai, Jian Cao, Xin Qian, Weidong Yang
School of Computer Science, Fudan University, Shanghai, P. R. China;
XAC Commercial Aircraft Co.,Ltd, Xi'an, P. R. China

Abstract: As most system requirements are written in natural language, a flexible and informal representation, they are likely to carry natu-ral defects such as ambiguity, incompleteness, and non-atomicity. Although researchers have put great effort into defect detection of natural language requirements, the work of building public datasets is still not as sufficient as those in other fields of NLP. In our work, we present PARD (Public-Available Requirement Defects) Dataset, a dataset of 148 publicly available requirement documents or specifications integrated from several sources, together with a pipeline model we proposed to transform documents into collections of requirement statements and labels of natural language defects. The model also shows the possibility of custom-izing the pipeline to satisfy the demand of specific industrial scenarios.

Detection of Virus Integration Sites in Tumor Genomes Using Deep Convolutional Neural Networks

Vibhuti Gupta, Emirrah Sanders
School of Applied Computational Sciences, Meharry Medical College, Nashville, Tennessee, USA;
School of Graduate Studies and Research, Meharry Medical College, Nashville, Tennessee, USA

Abstract: Pathogenic viruses are estimated to be responsible for 15% of all human cancers globally and pose significant threats to public health. Viruses integrate their genetic material to host genome and increases the risk of cancer promoting changes in it. To understand the molecular mechanisms of viral mediated cancers, a necessary step is to detect virus insertion sites in cancer genomes, however the challenges posed by exponentially increasing volume of tumor sequencing data and by accurate analysis of data needs to be addressed. In this paper, we propose a deep convolutional neural network (CNN) based framework for detecting virus integration sites in tumor genomes. Our contributions are twofold: (i) we design a novel approach by constructing image pairs from NGS data including reads from host and virus genome; (ii) we utilize one-hot encoded images with lower computational complexity to represent viral integration sites and leverage the power of twin Deep CNN networks to perform the detection; (iii) we integrate above techniques into an end-to-end framework to produce next-generation tool for NGS-based detection of virus integration. Finally, we point out related datasets for experiments and future research directions.

A Graph Convolutional Network with Recurrent Neural Network Structure for Heterogeneous Financial Fraud Data

Sicheng Shen
Mathematics and Economics Department, Oberlin College, Oberlin, Ohio, USA

Abstract: With the development of Internet financial technology, traditional fraud detection technology has been greatly challenged. The existing financial fraud detection methods usually come from transaction data, which generally have data imbalance and makes the classification and detection performance unable to be further improved. In this paper, a graph convolutional network with a recurrent neural network structure is proposed to extract and classify features that describe node associations and contextual associations by modeling heterogeneous data. The experimental results on the simulated dataset show that this structure can achieve better classification index evaluation results on the simulated datasets.

Population Density, Health, Education and Dwelling Quality Effects on Social Lag Status of Mexican Localities

Ana Sofia Landeros Basurto, Crhys Lorena Amelia Torres Bautista, Eitan Arakanchi Weil
School of Engineering and Science, Tecnologico de Monterrey, Monterrey, Mexico

Abstract: A study on the relationships between social lag, health, education, dwelling quality and population magnitude of Mexican localities was made using a multidimensional poverty measurement dataset. Using data mining techniques and comparing the performance of different classification models, Support Vector Machines (SVM) classification model was used to determine localities social lag status. Predictions were made and the results compared to the social lag status of the training set. This classification model suggests that illiteracy, dwelling conditions and basic education are the main determinants of social lag status in localities. The contribution of these results and model provides a deeper understanding of the social lag in Mexico seeking to increase social welfare by designing more efficient and adequate public policies.

Analyzing Operational Technology Data with Multiple Correspondence Analysis and Factorial Analysis of Mixed Data

Hunter Sylvester, David Schweitzer
USACE ERDC ITL CEAB R&D - US Army Corps of Engineer, Information Technology Laboratory (ITL), Engineer Research and Development Center, USA;
The U.S. Army Engineer Research and Development Center (ERDC), USA

Abstract: The increase of cyber-attacks affecting countries on a large scale, such as the Colonial Pipeline ransomware attack, have warranted and created an urgency for more research in Industrial Control System (ICS) and Supervisory Control and Data Acquisition (SCADA) network security. However, the lack of data pertaining to such systems due to its sensitive nature creates a bottle neck for progress. This paper begins an investigation into the fundamental data science of energy management system based operational technology (OT) data through the usage of multiple correspondence analysis (MCA) and factorial analysis of mixed data (FAMD). Both MCA and FAMD show consistent analysis in their results, and we demonstrate the ability to identify key factors in anomaly detection despite the data having very low variation.

Improved Sample Type Identification for Multi-Class Imbalanced Classification with Real-World Applications

Jiawen Kong, Wojtek Kowalczyk, Kees Jonker, Stefan Menzel, Thomas Baeck
Leiden University, Leiden, Netherlands

Abstract: Driven by studying the nature of imbalanced data, researchers proposed to consider different types of samples (safe, borderline, rare samples and outliers) in the minority class. The idea was first proposed and evaluated on binary imbalanced classification problems and then extended to multi-class scenarios. However, simply extending the identification rule in binary scenarios to multi-class scenarios results in several problems, for example, a higher percentage of unsafe samples in minority classes and a false identification of outliers. In this paper, we first show the drawbacks when extending this idea from binary to multi-class scenarios. Then, we propose a new identification rule for multi-class scenarios. In our experiments, we consider oversampling different types of samples before performing classification, where oversampling is a data-level approach to deal with the imbalance in the datasets. Experimental results on benchmark datasets indicate that the proposed rule can decrease the probability of false identification and improve the classification performance on minority class(es) on average by 7.4%. In addition, we apply our proposed rule on surface inspection data from the steel industry and confirm its effectiveness and potential usefulness in real-world applications.

Deep Latent Similarity Model for Online Data Categorization

Shuanglu Dai, Hong Man
Stevens Institute of Technology, New Jersey, USA

Abstract: A deep latent similarity model (DLSM) is introduced as an unsupervised online data categorization method that is able to continuously characterize data series with countably infinite samples and categories. In this method, a distance measure based on a deep metric network is introduced to capture non-linear data similarities. An affinity propagation framework is then adopted, in which a similarity function is introduced to characterize latent exemplars. The proposed method attempts to infer exemplars as ``pseudo categories'' from a knowledge data set, using an cumulatively learned similarity measure from data series in an online fashion. Separability conditions are examined to further constrain the similarity function in order to generate valid exemplar assignments. Finally, a probabilistic graph is proposed to model the joint distribution of the network distance, similarity function and latent exemplars. Competitive results on online clustering, unsupervised image categorization and extreme clustering demonstrate the effectiveness of the proposed model.

High-Level Synthesis Parallelization and Optimization of Vectorized Self-Organizing Maps

Omar Rivera Morales, Lutz Hamel
University of Rhode Island, Rhode Island, USA

Abstract: The nature of the Self-Organized Maps (SOM) requires a constant improvement of performance to address the increasing complexity of datasets. These demands have led to high-performance algorithms that run in hardware accelerators such as Graphical Processing Units (GPU) and Field Programmable Gate Array (FPGA). This work introduces a novel High-Level Synthesis (HLS) FPGA implementation for the vectorized SOM algorithm. The proposed algorithm is implemented using HLS parallelization and design optimization techniques available on the Xilinx Alveo FPGA Accelerator Card. This paper introduces the HLS-based algorithm and discusses the pipelining, unrolling, systolic array matrix reduction, and memory transformation techniques to improve the VSOM algorithm performance. Our HLS-VSOM experimental results show a significant performance increase over SOM CPU and parallel GPU variants.

Covi-Care: Secure Covid-19 Vitals Diagnosis and Disease Identification Application

Anurag Gorkar, Pranav Patil, Yogesh Kulkarni, Sarang Joshi
Pune Institute of Computer Technology, India

Abstract: Covid-19 and its variations continue to infect thousands of people worldwide, and even persons who have been properly vaccinated are being infected again. This study aims to provide a precise early diagnosis for Covid-19 in an easy-to-use mobile application that includes a Convolutional Neural Net (CNN) for identifying Covid patients using a chest CT-scan with a 76% accuracy, as well as an embedded device for capturing critical vitals and using them to build a Covid-19 susceptibility score. Furthermore, the application employs an Artificial Neural Net (ANN) to serve as a mental health assessor for professionals, predicting if the employee should seek professional therapy with an accuracy of 81%. This application also contains a cough sound classifier, which predicts if a person is infected with Covid-19 using cough audio spectrogram generated feature space and has a 96% accuracy. Finally, for Covid-19 patients, this application integrates a U-net-based infection mask generator that constructs an infection mask around the infected region in the CT-scan with a dice coefficient of 0.84. A video demonstration of this work is provided in https://youtu.be/qxpOh35byF8 .

Improving Debt Collection Forecasting Models using Gradient Boosting Machines and Neural networks based on the Impact of the COVID-19 Pandemic

Teo Curcic, Cristian Rodriguez Rivero, Allard de Winter
University of Amsterdam, Amsterdam, Netherlands

Abstract: The COVID-19 pandemic that started in 2020 had a negative influence on many businesses worldwide, leading to financial uncertainties within different business areas. As a result, many companies were unable to pay their monthly electricity usage on time to energy suppliers. This created a continuous loop of accumulating financial liabilities for many companies that rely on services from others, while not being able to pay timely. To remain competitive during such difficult times with focus on limiting financial losses, energy suppliers rely on the power of data to optimize business to business customer processes. One method is by using machine learning to improve current debtor models that predict the probability of default. This paper presents an analysis of customer payment behaviour data using advanced machine learning approaches, such as XGBoost, LightGBM and LSTM, to anticipate the probability of default as a main business delivery. Results of this work show that LSTM outperforms the current debtors models as it utilizes the persisting trends changes in the second half of 2020 and focuses on the most significant customer segmentation groups such as high/low monthly payers and high/low credit risks scores.

Unsupervised Twitter Sentiment Analysis using Emoji and Context-Based Analysis

Aksht Jain, Apoorva NaN, Aiman Siddiqua, Trasha Gupta
Delhi Technological University, India

Abstract: Twitter is one of the largest social media platform and observed a large volume of tweets from cricket fans worldwide during the T20 Cricket World Cup 2021. People used the platform to express their emotions and opinions throughout the series. Sentiment analysis is an effective tool for gauging public opinion on various subjects, for instance, surveying the impact of a newly released product, predicting election results etc. In this paper, we create a dataset of tweets collected during the knockout stages of the World Cup. We also propose an unsupervised framework using Emojis and Context based analysis to identify the corresponding emotions. The proposed method outperformed two existing standard sentiment analysis libraries, TextBlob and VADER. We validated the findings by applying the model to two other publicly available annotated datasets, Airline Dataset and IMDB Movie Reviews Dataset. The results from these datasets prove that the proposed framework is a viable method for sentiment analysis.

Accident Severity Prediction: A Comparative Case Study

Nassim Sohaee
University of North Texas, Texas, USA

Abstract: N/A

Cointegration of SARS-CoV-2 Transmission with Weather Conditions and Mobility during the First Year of the COVID-19 Pandemic in the United States

Hong Qin, Syed Tareq, William Torres, Megan Doman, Cleo Falvey, Jamaree Moore,
Meng Tsai, Yingfeng Wang, Azad Hossain, Mengjun Xie, Li Yang
University of Tennessee at Chattanooga, Tennessee, USA

Abstract: Correlation between weather and the transmission of SARS-CoV-2 may suggest its seasonality. We examined the cointegration of virus transmission with daily temperature, dewpoint, and confounding factors of mobility measurements during the first year of the pandemic in the United States. We examined the cointegration of the effective reproductive rate, Rt, of the virus with the dewpoint at two meters, the temperature at two meters, Apple driving mobility, and Google workplace mobility measurements. Dewpoint and Apple driving mobility are the best factors to cointegrate with Rt. The optimal lag is two days for cointegration between Rt and weather variables, and three days for Rt and mobility. We observed clusters of states that share similar cointegration results, suggesting regional patterns. Our results support the correlation of weather with the spread of SARS-CoV-2 and its potential seasonality.

Using Machine Learning to Enhance Win Ratio for College Ice Hockey Teams

Sadixa Sanjel, Zelalem Denekew, Ahmed Sadek, Naseef Mansoor
Minnesota State University, Mankato, Minnesota, USA

Abstract: N/A

Understanding the Health Consequences of Air Quality using Machine Learning

Kevin Williams, Mohammad Pourhomayoun
California State University, Los Angeles, California, USA

Abstract: Elevated levels of pollutants are known to have a detrimental effect on one's respiratory health and depending on the season, weather can inflate pollution levels causing more harm to those with respiratory ailments. This study involves using averaged weather data and air pollutant levels to predict the amount of daily emergency department visits for asthmatic patients in Los Angeles County. We employed various machine learning algorithms such as Random Forest (RF), Multi-Layer Perceptron (MLP), XGBoost and Support Vector Regressor (SVR) with the best results coming from Random Forest, RMSE of 20.30, and the worst results from MLP, RMSE of 33.09.

Product Quantization for Surface Soil Similarity

Haley Dozier, Althea Henslee, Ashley Abraham, Andrew Strelzoff, Mark Chappell
US Army Engineer Research and Development Center, USA

Abstract: The use of machine learning (ML) techniques has allowed rapid advancements in many scientific and engineering fields. One of these problems is that of surface soil taxonomy, a research area previously hindered by the reliance on human-derived classifications, which are mostly dependent on dividing a dataset based on historical understandings of that data rather than data-driven, statistically observable similarities. Using a ML-based taxonomy allows soil researchers to move beyond the limitations of human visualization and create classifications of high-dimension datasets with a much higher level of specificity than possible with hand-drawn taxonomies. Furthermore, this pipeline allows for the possibility of producing both highly accurate and flexible soil taxonomies with classes built to fit a specific application. The machine learning pipeline outlined in this work combines product quantization with the systematic evaluation of parameters and output to get the best available results, rather than accepting sub-optimal results by using either default settings or best guess settings.

Large-Scale Data Parallelization of Product Quantization and Inverted Indexing Using Dask

Ashley Abraham, Andrew Strelzoff, Althea Henslee, Haley Dozier, Mark Chappell
US Army Engineer Research and Development Center, USA

Abstract: Large-scale Nearest Neighbor (NN) search, though widely utilized in the similarity search field, remains challenged by the computational limitations inherent in processing large scale data. In an effort to decrease the computational expense needed, Approximate Nearest Neighbor (ANN) search is often used in applications that do not require the exact similarity search, but instead can rely on an approximation. Product Quantization (PQ) is a memory-efficient ANN effective for clustering all sizes of datasets. Clustering large-scale, high dimensional data requires a heavy computational expense, in both memory-cost and execution time. This work focuses on a unique way to divide and conquer the large scale data in Python using PQ, Inverted Indexing and Dask, combining the results without compromising the accuracy and reducing computational requirements to the level required when using medium-scale data.

Machine Learning Framework for Traffic Congestion and Equity

Althea Henslee, Haley Dozier, Ashley Abraham, Andrew Strelzoff, Alexandre Ligo, Jules White
U.S. Army Engineer Research and Development Center, USA

Abstract: Consistent heavy traffic congestion places an undue burden on areas of higher poverty. Traffic congestion is also in many urban areas largely created by outside forces, i.e. commuters who do not face the same level of consequence as those living in the affected area. Reinforcement learning agents are particularly well-placed to address this equity issue, though an environment based on real world traffic data is needed first. This work provides an environment designed using historical data from the last decade from the California Department of Transportation Performance Management System. This environment will allow future reinforcement learning models to identify areas with consistently high congestion, and devise new routes to alleviate congestion in the greater Los Angeles area while lessening the impact on high poverty communities.

Robust Generalized High-Order SVD for Images with Outliers and Correlated Noises

Jionghua Jin, Yaser Zerehsaz
University of Michigan, USA

Abstract: The SVD (singular vector decomposition) method is popularly used for variation pattern analysis in the applications of system monitoring and faulty diagnosis. With the increasing use of image-based machine vision systems, the generalized high-order SVM (GHOSVD) method has been developed to handle high-dimensional images with correlated noises. However, the performance of GHOSVD is sensitively affected by inevitable outliers due to image sensor errors. Instead of requiring a separate data cleaning step, this paper proposes a new Robust GHOSVD (RGHOSVD) method that can simultaneously handle outliers and spatiotemporally correlated noises in images. The performance of the method is evaluated using simulation studies.

Learning Sign Language Representation using CNN-LSTM, 3DCNN, CNN.RNN.LSTM and CCN-TD

Koffka Khan, Nikita Louison, Wayne Goodridge
University of the West Indies, West Indies, Jamaica

Abstract: Existing Sign Language Learning applications focus on the demonstration of the sign in the hope that the student will copy a sign correctly. In these cases, only a teacher can confirm that the sign was completed correctly, by reviewing a video captured manually. Sign Language Translation is a widely explored field in visual recognition. This paper seeks to explore the algorithms that will allow for real-time, video sign translation, and grading of sign language accuracy for new sign language users. This required algorithms capable of recognizing and processing spatial and temporal features. The aim of this paper is to evaluate and identify the best neural network algorithm that can facilitate a sign language tuition system of this nature. Modern popular algorithms including CNN and 3DCNN are compared on a dataset not yet explored, Trinidad and Tobago Sign Language as well as an American Sign Language dataset. The 3DCNN algorithm was found to be the best performing neural network algorithm from these systems with 91% accuracy in the TTSL dataset and 83% accuracy in the ASL dataset.

Three-Dimensional Comparative Analysis of Large Populations to Determine Data Relationships: An Alternative to Statistics

Vera Dolan
VFD Consulting, Inc., Las Vegas, Nevada, USA

Abstract: An alternative to the use of statistics leads to precision measurements of any large population. For large populations of interest, three-dimensional visualization of two descriptive variables graphed respectively against an outcome variable as well as their population percent distribution is a robust analytical measurement tool when the two three-dimensional graphs are compared with each other. This alternative analytical measurement tool has provided solutions to problems that have been insoluble using a traditional statistical approach. An example of such a problem involves the acknowledged inaccurate current status of laboratory test reference intervals that are not specific by age and sex. New laboratory reference interval lookup charts that more accurately and precisely define what laboratory test results are normal or abnormal by age and sex can be built through use of this alternative analytical method. Other examples of applications that can benefit from such precise measurements include electric car batteries, solar panels, agricultural crop yields, drinking water quality and defense capabilities.

An Ensemble Learning Augmentation Method for Concept Drift Detection over Streaming Data

Ali Alizadeh Mansouri, Abbas Javadtalab, Nematollaah Shiri
Concordia University, Canada

Abstract: We study concept drift detection over streaming data, and investigate suitable solutions based on the natures and rates of the drift. We propose an online incremental streaming ensemble learning method that augments drift detection. For this, we study different techniques based on bagging, boosting, and verification with the main classifier, each of which is applicable to different rates of drift. Using bootstrapping on top of these techniques, we managed to reduce the overall variance of the detection algorithm, hence improved the performance of concept drift prediction as well as the accuracy of the main classifier. The results of our numerous experiments using synthetic and real-world datasets indicate significant improvement of F1 score compared to existing methods with fewer re-modelings of the main classifier. We obtained three times or more improvement of detection F1 score over the state-of-the-art method with 92% reduction in computation time, and ten times improvement over a persistent forecast classifier with less than 13% of the number of re-modelings on datasets with abrupt drifts. Our results yield guidelines for deciding suitable effective solutions for concept drift detection, by providing a balance between prediction quality and computational cost for different datasets and applications.

DeepPoint: A Hybrid of Deep Neural Network and Self-Exciting Point Process for Street Crime Prediction in New York Crime Data

B Shanker Jaiswal, B Chandra, Kolin Paul
Indian Institute of Technology (IIT), Delhi, India

Abstract: This paper proposes a novel crime prediction algorithm DeepPoint- a hybrid of deep learning and point process. It aims to improve the prediction of high-volume street crimes. In the design of this learning pipeline, census attributes have also been factored in as features in addition to space-time features in the crime dataset. Furthermore, the conditional background and triggering intensities of crime have also been used as features, which are derived from the modified self-exciting point process. The most important features are selected and used as input to deep neural and convolutional neural networks to enhance the accuracy of crime prediction. The NYPD's crime dataset and New York City's census dataset have been used to illustrate the effectiveness of the algorithm.

Forecasting Sensor-data in Smart Agriculture with Temporal Fusion Transformers

Boje Deforce, Bart Baesens, Jan Diels, Estefania Serral Asensio
Katholieke Universiteit Leuven, KU, Belgium

Abstract: Smart agriculture has grown in significance over the past years. Agriculture comes with certain inherent complexities (e.g. field variability, etc.). However, the advent of big data and efficient learning algorithms has enabled the field to model and predict seemingly complex systems. Here, we focus on five-day-ahead quantile forecasting of soil moisture captured by in-field sensors. Forecasting soil moisture is not trivial since it depends on many exogenous factors. We suggest the use of Temporal Fusion Transformers to build a global interpretable forecasting model that can include static (e.g. field characteristics) and temporal (both future and historic) features. Existing methods typically build local models (e.g. one per field) thereby potentially missing the opportunity for the model to learn from other fields too. We demonstrate this method on a challenging dataset coming from three different pear orchards in Belgium. Results suggest that Temporal Fusion Transformers are helpful in accurately forecasting soil moisture.

High-Resolution Spatiotemporal PM2.5 Prediction with Deep Convolutional LSTM using Atmospheric and Ground-level Data

Pratyush Muthukumar, Dawn Comer, Jeanne Holm, Mohammad Pourhomayoun,
Shaurya Pathak, Kabir Nagrecha, Navid Amini
California State University, Los Angeles, California, USA

Abstract: N/A

Cluster Analysis on Operational Technology Data Using K-Modes

Hunter Sylvester, David Schweitzer
The U.S. Army Engineer Research and Development Center (ERDC),
U.S. Army Corps of Engineers, USA

Abstract: Operational technology (OT) data differs from traditional IT data in that it is more constrained and not understood from an abstract perspective. There is a need to understand OT data more abstractly due to OT anomaly detection's weakness in relying on understanding the system, which is usually well-defined, rather than the data, which leads to poor generalizability. This paper investigates the benefits of clustering packet captured data through facilities network and how such clustering can assist in identifying previously unknown relationships in OT data that can lead to improved anomaly detection.

Examining Operational Technology Data Abstractly Using Principal Component Analysis

Cheronika Manyfield-Donald, David Schweitzer
The U.S. Army Engineer Research and Development Center (ERDC),
U.S. Army Corps of Engineers, USA

Abstract: In this paper we examine Operational Technology (OT) data using Principal Com-ponent Analysis to detect features in the dataset that can be clustered due to high correlation for future predictive analysis. Our results show, Principal Component Analysis, as a beginning step to analyzing OT data, can successfully be used to re-duce high dimensionality in the data to assist with extracting correlated features so predictive analysis can be performed on the features that most impact the system.

Application of Linear Regression and Random Forest Models in Predicting Bus On-Time Performance

William Noonan, Andrew Trombly, Jeong Yang
Department of Computing and Cyber Security, Texas A&M University-San Antonio, Texas, USA

Abstract: Machine Learning is a technique used to automatically extract valuable patterns from data. Several types of algorithms have been used to create successful Machine Learning models. The purpose of this paper is to discuss some supervised and unsupervised machine learning algorithms and to compare the performance of two algorithms based on bus transit data from the 2020 San Antonio Datathon that was provided by VIA Metropolitan Transit. The performance of a Linear Regression model and a Random Forest model was compared in predicting bus on-time performance using the panda's data science library and the Scikit-learn (sklearn) machine learning library. Their performance was evaluated by computing the root-mean-square error (RMSE) between their predictions and actual values. The Linear Regression model yielded an RMSE very close to zero and significantly outperformed the Random Forest Regression model, which yielded an RMSE of about 3. This is likely because of the strong linear trends that exist in the underlying datasets as well as the strong possibility that the independent variables in the training data were not truly independent. This analysis illustrates the importance of exploring the trends and relationships in data before selecting a machine learning model.

Bimodal Classification with CNN and Sequence Models

Wookey Lee, Keywoong Bae, Mukhiddin Toshpulatov, Suan Lee
Inha University, Republic of Korea

Abstract: N/A

When to Publish on Twitter in Mexico

Jose Manuel Tapia Avitia, Carlos Alonzo Lopez-Castaneda
Tecnologico de Monterrey, Mexico

Abstract: For most of the people, the use of social media is included in their daily routines. Therefore, investing in marketing strategies that target social media usage is essential for any business. These media offer great potential to advertisers to reach their consumers. Nevertheless, making a publication in the early morning on Mondays is not the same in an early afternoon on Wednesday. People who prefer to talk about technology could be active at a different time than people that used to talk about food. The marketer could consider the context to take advantage of this behavior to make a successful marketing campaign. This research work retrieves tweets from March 5 to March 25 related to specific topics in Mexico. We use such tweets to identify at which time of the day there is more activity around each topic. In addition to such an analysis, we consider the category of sentimental analysis using a pre-trained model. This information allows us to provide a more fruitful and valuable report for the marketer.

Blockchain-based Decentralized Identity Management for Privacy-preserving over 6G Smart Healthcare

Sejong Lee, Jaehyeon Kim, Yongseok Kwon, Yushin Kim, Sunghyun Cho
Hanyang University, Republic of Korea

Abstract: N/A

Automatic Generation of Class Diagrams using Deep Learning and Candidate Recall

Eric Xu
Zhuhai Fudan Innovation Institute, P. R. China

Abstract: Requirement analysis and modeling are major processes in requirements engineering that are both complex and time-consuming, and require substantial expertise. In practice, the UML class diagram is widely used and essential for requirement modeling and application of object-oriented programming paradigm. Researchers have been studying on the automation of generating high quality class diagram to facilitate the process and reduce the cost. Previous methods were mainly focused on rule-based approach which lacks flexibility and scalability. Recently, many studies on leveraging deep learning models during the generation process have emerged. We propose a new method with a gain of 0.9 F1 score and improves the quality of generated UML class diagram up to 7.6% on our test dataset.

The International Conference on Emergent Quantum Technologies (ICEQT'22: July 25-28, 2022, USA)
https://baylor.ai/iceqt/
https://american-cse.org/csce2022/

An Overview of Quantum Probability Support for Robotic Object Tracking

Shahram Payandeh
Networked Robotics and Sensing Laboratory, School of Engineering Science,
Simon Fraser University, British Columbia, Canada

Abstract: Tracking movements and activities of objects such as people, wildlife, or autonomous intruder robots through various sensing modalities has many applications. These applications can range from health-related monitoring issues to animal migration and to security and surveillance systems. Stochastic-based tracking approaches such as particle filter has been proposed and applied in the past with a high degree of success. One of the challenges of such approaches is the development and representation of a model where various levels of uncertainties in measurements/observations and movements can be used to interpret the outcome of events in a sensor network. In this paper, we present an overview of the quantum probability framework as a complementary presentation to the classical methods in the context of object tracking. The notion of classical conditional probability is represented as a mapping(projection) between various subspaces. Through a simple tracking example, we describe the proposed framework and its possible interpretation.

An Implementation of the Grover Unstructured Search Algorithm

Hugo J. Delgado-Marti, Alfredo Cruz
Polytechnic University of Puerto Rico, Puerto Rico

Abstract: Grover algorithm brings polynomial advantage in search complexity that can be ad-vantageous in artificial intelligence. We propose an implementation of a Grover search for an inversion matrix using amplitude amplification. A known biased distribution that has been modified by a marked element can be recovered if the inversion matrix from an expected distribution can be found. The controlled-swap test is used to find such a solution if the system returns to a uniform probability distribution after suc-cessive multiplications and inversions. The algorithm can be improved using Phase Oracles that do not require ancilla qubits, but our implementation failed to work in a real quantum computer. We propose that such systems could be used as part of a search algorithm where an element introduces a bias to an expected distribution.

Kernels and Quantum Machine Learning

Bikram Khanal
School of Engineering and Computer Science, Department of Computer Science,
Baylor University, Texas, USA

Abstract: Support Vector Machine (SVM) and Kernel Method (KM) are used widely for classification and regression in learning from data. Kernels are positive definite functions that map the data into higher (possibly infinite) dimensions. Generally, SVMs implement kernel methods as a subroutine that maps the non-linear data to a higher dimensional where it becomes linearly separable. SVM draws a linear decision boundary between classes of data points in this feature space. This paper reviews kernels and kernel methods from a classical machine learning perspective and their possible implementation in quantum machine learning. We start with the basis of kernels, including Hilbert space and Reproducing Kernel Hilbert Space, Mercer's condition, and prove three widely used kernels validity satisfying Mercer's condition. We review two different approaches of quantum machine learning, parameterized quantum circuit and kernel-based training, and discuss the potential advantage of one over another. This paper can facilitate the readers' getting started with kernel theory and quantum machine learning.

A Brief Summary of Quantum Computing Research

Curtis Busby-Earle
Department of Computing, The University of the West Indies, Mona, Jamaica

Abstract: Quantum computing is an emerging field. There is much ongoing research activity related to many of the aspects involved with the realization of quantum computers, quantum networks, their commercialization and implementation; from qubits with long coherence at room temperatures to government led initiatives to develop post-quantum computing cybersecurity systems. In this paper we summarize some of the ongoing research activity categorized according to the logical levels of a quantum computer. In particular, attention is focused on the higher, more abstract layers and also provide a few open areas of research that have been identified.

Quantumized Graph Cuts in Portfolio Construction and Asset Selection

W. Bernard Lee, Anthony G. Constantinides
HedgeSPA Limited, Singapore;
School of Accounting and Finance, Hong Kong Polytechnic University, Hong Kong;
Financial Signals Processing Laboratory, Imperial College London, London, United Kingdom (UK)

Abstract: This paper is concerned with two fundamental problems in investment science, namely (a) the construction of a portfolio by segregating target assets into sectors representing a typical market index and (b) the selection of target assets from each of those sectors. Such solutions may be applied, for example, to construct a portfolio of 50 assets (say) that aims to outperform the S&P 500 index by selecting the most promising performer in each applicable sector or subsector. The formulation of this investment objective is non-trivial. This paper addresses the above problems and offers important observations and potential solutions.

Time-Efficient Quantum-to-Classical Data Decoding

Naveed Mahmud, Mingyoung Joshua Jeng, Md. Alvir Islam Nobel, Manu Chaudhary,
SM Ishraq-Ul Islam, David Levy, Esam El-Araby
Electrical Engineering and Computer Science, University of Kansas, Lawrence, Kansas, USA

Abstract: A critical challenge for Noisy Intermediate-Scale Quantum (NISQ) devices is achieving time-efficient quantum state measurement or readout. Generally in quantum applications, data is encoded as the amplitudes of a superimposed quantum state. To extract information from the quantum state, the quantum circuit is sampled repeatedly, which incurs significant overhead in the system execution time. In this paper, time-efficient methods for decoding information from quantum states are proposed and evaluated. The process of extracting classical data from the quantum domain is termed in this work as quantum-to-classical (Q2C) data decoding. We propose a novel Q2C approach based on time-efficient sampling of quantum states using multi-level decomposable Quantum Wavelet Transform (QWT). Experimental evaluations of the proposed Q2C method are performed on a state-of-the-art quantum computing platform from IBM Quantum. Measurements of circuit execution time and circuit depth are obtained. Experimental results are consistent with our theoretical expectations and we present a quantitative comparison with existing techniques that confirm the efficiency of our proposed approach.

A Quantum Algorithm for Proof-of-Work

Fabiola Canete-Leyva, Carlos A. Perez-Delgado, Salvador E. Venegas-Andraca
Faculty of Physical and Mathematical Sciences, Benemerita Universidad Autonoma de Puebla, Mexico;
School of Computing, University of Kent, Canterbury, England, UK;
Computer Science Department, Tecnologico de Monterrey Campus Santa Fe, Mexico

Abstract: Quantum computers provide efficiency advantages over their classical counterparts, for various computational problems. One such problem is "mining" cryptocurrencies. Mining is the process, in Proof-of- Work (PoW) cryptocurrencies such as Bitcoin, by which individuals are rewarded for solving computationally hard problems. It has been stated that quantum miners could have a computational advantage over classical ones. In this work, we develop an explicit quantum algorithm for performing proof-of-work. We then provide an in-depth analysis of this algorithm's precise advantage over classical miners.

Securing IPsec Communications with Quantum Key Distribution

Leo Truksans, Elina Kalnina, Andris Lapins, Edgars Rencis, Juris Viksna
Institute of Mathematics and Computer Science, University of Latvia, Riga, Latvia

Abstract: We present a method for securing IPsec communications with keys generated by Quantum Key Distribution (QKD) devices. With commercial QKD devices becoming increasingly available their integration into the existing communication networks has become an important research and practical problem, with IPsec being a very suitable choice of cryptographic protocol for which to provide such integration. Compared to other work, we do not attempt to propose a modified protocol inherently incorporating QKD but to provide IPsec access to QKD-generated keys via an additional software layer. Although such an approach has limited capacity to comply with stringent security and reliability requirements, the advantage is the ease of implementation that allows for testing the performance of yet emerging technology within the framework of existing communication infrastructure.

Knowledge Navigated Quantum-Inspired Tabu Search Algorithm for Reversible Circuit Synthesis

Shu-Yu Kuo, Yao-Hsin Chou, Shan-Jung Hou, Yu-Yuan Chen, Hsing-Yu Hsu, Yu Chen
National Chung Hsing University, Taiwan, R.O.C.;
National Chi Nan University, Taiwan, R.O.C.

Abstract: Quantum computing has received remarkable attention and has great potential to be the next computing paradigm. All operations in quantum circuits are reversible. Therefore, this paper proposes a novel reversible circuit synthesis method based on the knowledge navigated quantum-inspired tabu search (KNQTS) algorithm using Toffoli and Fredkin gates, to facilitate the development of quantum technologies. KNQTS has powerful search abilities and uses a knowledge navigation strategy to detect its convergence speed and adaptively adjust its parameters to strengthen both exploration and exploitation. The experimental results show that KNQTS has promise and can use fewer costs to achieve the same function compared to traditional methods.

The 23rd International Conference on Internet Computing & IoT (ICOMP'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-ICOMP
https://american-cse.org/csce2022/

External Key Brokerage for Security of Cloud Data

Gil Mazor, Asad Ali
Thales, Digital Identity & Security, Petach Tikva, Israel;
Thales, Digital Identity & Security, Austin, Texas, USA

Abstract: As businesses migrate their data and services to the cloud, there is an increasing need to keep this data secure. Cloud Service Providers (CSP) do this by encrypting the data and managing cryptographic keys themselves. However, to address regulatory compliance and achieve greater security, customers are beginning to demand control over the protection of their data. This paper explains a decentralized model of key management where control of cryptographic keys can be brokered through third parties and handed back to owners of the data. This separation of duty between data storage providers and key management specialist can create a sustainable and secure framework for CSP customers. Data stored in a public cloud can be as secure and protected as data in traditional enterprise on-premises data centers.

Online Social Network Influencers

Mehran Asadi, Afrand Agah, Clayton Amondo
Lincoln University, USA;
West Chester University, USA

Abstract: N/A

WiFi CSI Subcarrier Selection Methods for Sensing Applications

Jesus A. Armenta-Garcia, Felix F. Gonzalez-Navarro, Jesus Caro-Gutierrez
Engineering Institute, Universidad Autonoma de Baja California, Mexicali, Mexico

Abstract: Subcarrier selection is a step in WiFi CSI-based sensing applications that consists of selecting subcarriers from a WiFi signal that contain relevant information for a sensing task according to signal characteristics. This step is often omitted or not considered in the performance evaluation when using different subcarrier selection methods in the literature. In this paper, three different common subcarrier selection methods for WiFi CSI-based sensing applications are compared to observe how they improve a breathing rate estimation and classification task. Experimental results show a RMSE decreasing from 1.511 to 1.41 by using only 20 relevant subcarriers out of 90.

Modbus to IEC61850 Converter: Integrating Configuration Tool with Run-time

Abdelaziz Said Abdelaziz, Hassan Bedour, Hoda K. Mohamed
Faculty of Engineering, Computers and Systems Department, Ain Shams University, Cairo, Egypt

Abstract: In many power substations, we need to transfer Modbus measurements like volts, currents, frequencies, and powers from the high or medium voltage levels to the station level as IEC61850 standard, so we need to automate and integrate Modbus with IEC61850 standard into one entity to send alarm points to the IEC61850 client automatically in case of a defect in measurements by the Modbus slaves. The paper discusses integrating a configuration tool for physical devices with a run-time tool into the entity to perform Modbus to MMS mapping and create the dedicated IED configuration file. Configuration Tool dispenses with creating the preconfigured mapping associated file, device description file, IED configuration file by an external tool. Run-time tool dispenses with the database for mapping for each Modbus response.

An Evaluation of Machine Learning and Deep Learning Approach on Ekman Sentiment Classification

Abdou Aziz Ledoux Ngou Njikam, Elie Fute Tagne, Lionel Landry Sop Deffo
Nanjing University of Science and Technology, Nanjing, P. R. China;
University of Buea, Cameroon;
University of Dschang, Dschng, Cameroon

Abstract: The evolution of natural language processing during the last years has been phenomenal. Its applications are infinite and the society as we've noticed can no longer do without it. Works in natural language processing have reached a point where they have been able to successfully collide with those in psychology. The outcomes of this are extraordinary. One of which we focus our interest is the basic emotion elaborated by the psychologist Ekman who classifies emotions into six basic states namely anger, disgust, fear, joy, sadness and surprise. This work focuses on the evaluation of machine learning and deep learning approaches based on the Ekman's emotion classification. It first investigates and evaluates on four machine learning algorithms which are random forest, support vector machine, naive Bayes, and logistic regression. Then two deep learning models are tested as well: LSTM (Long Short-Term Memory networks) and BERT (bi-directional encoder representation). The dataset used is the Emotion Lines dataset. The study finishes by a comparative study of the different approaches.

IoT, from Collection, to Data Processing, Edge Computing, to Smart Applications

Hassan Rajaei
Computer Science Department, Bowling Green State University, Bowling Green, Ohio, USA

Abstract: Innovations in Internet of Things (IoT) are creating new waves of smart applications such as cars, cities, homes, and buildings as primary targets. The IoT ecosystem continuously collects data from multiple sources of the targeted applications, monitors the connected devices, transports the data, process at designated datacenters, and provide feedbacks to their smart applications. This paper focuses on the transport and data processing of IoT. We use 5G and its Edge Computing capabilities for data processing and argue that combination of 5G and Edge Computing deliver two major advantages for the IoT ecosystem: fast and broadband transportation, and speedy data processing at the edge. Adding analytics and AI in the processing loop, speedy intelligent feedbacks can be supplied for the smart applications and their end-users. Keywords: IoT (Internet of Things), Edge & Cloud Computing, Smart Applications, 5G Wireless Network.

The Implementation of the IoT Smart Home Devices' Security, Privacy, and Firmware Labeling System

Naif Waheb Rajkhan, Jia Song
Computer Science Department, University of Idaho, Moscow, Idaho, USA

Abstract: Smart homes are one of the great implementations of the Internet of Things (IoT). IoT smart home is an automation system that utilizes IoT devices that perform different functions. It provides powerful autonomous functions for homeowners to simplify and enhance the performance of connected devices. The IoT smart home device is a device that is connected to the internet to ease communication and perform actions inside the smart home environment. Since smart home devices collect and send data about the home environment and the homeowner over the internet, they could be an easy target for hackers if a device is not designed and equipped with the best privacy and security standards. Such attacks would harm the IoT smart home system, and gain access to homeowners' personal information and private environment data. Thus, devices with weak security and privacy standards should be avoided. However, a customer does not know whether an IoT device has poor security and privacy standards when purchasing a device. To help with that, we proposed an IoT smart home devices' security, privacy, and firmware labeling system. The labeling system is designed to be shown on IoT smart home devices' packaging and therefore provides additional information to the customers before making the purchase. This paper discusses such labels' implementation procedures and advantages in the Saudi Arabian IoT smart home devices market.

Cybersecurity from an Ethical Viewpoint

Isaac Mayo, Kathrine Stalemo, Kylie Hoff, Marideissa Dizon, Matthew Schanilec, Mohammed Mahmoud
Department of Computer Science & Information Systems, Minnesota State University, Moorhead, Minnesota, USA

Abstract: As Information Technology (IT) becomes increasingly ubiquitous in human society, the need for robust, resilient, and effective cybersecurity grows proportionately. Cybersecurity is defined by IBM as the practice of protecting critical systems and sensitive information from digital attacks. Such information is key to protect as it has implications socially, economically and politically. Just as cybersecurity is necessary for the health and well-being of information technology, ethics are necessary for the healthy application of cybersecurity. In this paper, we will review the current topics and trends in the domain of cybersecurity ethics. Applied ethical scenarios, such as ransomware attacks and other forms of computer crime, which affect both public and private sectors, will also be discussed. These ethical scenarios give rise to myriad ethical dilemmas for affected parties.

Reusable Test-Ready Models of Smart Home Systems

Afnan Albahli, Anneliese Andrews, Ahmed Al Haddad
University of Denver, Denver, Colorado, USA;
Princess Nourah Bint Abdul Rahman University, Riyadh, Saudi Arabia;
United Technology (U-Tech), Riyadh, Saudi Arabia

Abstract: The Internet of Things (IoT) has become one of the most important platforms in the last decade. The number of devices connected to the Internet has been increasing at a rapid rate, and is projected to continue. This raises a significant challenge for software testers. As a result, new testing methods must be explored to ensure the quality and effectiveness of such systems. This paper proposes reusable test-ready models of smart home systems. We proposed Extended Finite State Machines (EFSMs) and Communicating Extended Finite State Machines (CEFSMs) as test-ready models for device components and devices. We employed FSMApp to model the Mobile Application controlling the smart home system.

Sequential Recommender System for IoT Ontologies based on Graph and Semantic Embedding

Yuexian Li, Chuantao Yin, Xi Wang, Si Wu, Lan Wang
Sino-French Engineer School, Beihang University, Beijing, P. R. China;
Orange R&D Beijing Company Limited, Orange Company Limited, Beijing, P. R. China

Abstract: In the field of the Internet of Thing, developers often need to build avatars based on ontologies, how to recommend ontologies based on their created avatars is a meaningful study. However, because of lacking historical user-item data, it will be difficult to adopt conventional sequential recommender system in this task. In order to achieve good results in this situation, this paper proposes a sequential recommender system based on graph embedding and semantic embedding. We also use exponentially weighted average method to get user sequence embedding and sort recommendation results based on similarity matching algorithm. The experimental results show that the recommendation method proposed in this paper does not need a large amount of historical user-item data for training, and can obtain diversified recommendation results.

Stable Operation of Redundant Traffic Reduction System with Packet Cache by Relay Point Designation

Hiroto Hoshiba, Kenji Ichijo, Akiko Narita
Graduate School of Science and Technology, Hirosaki University, Aomori, Japan

Abstract: We often find the same data transferred through the same route repeatedly on the internet. Such redundancy should be repressed for efficient use of network re-sources. We have developed the traffic reduction (TR) node to reduce redundancy of concurrent transmission on TCP/IP network. One encoder TR node and at least one decoder TR node work cooperatively using synchronized packet cache. It is desirable for the TR nodes to configure encoder-decoder pair autonomously for effective use and easy management of the TR nodes. In this paper, we present a method by which the encoder TR node designate the decoder node as a pass point to offer stable service using newly defined IP option.

IoT Avatar Retrieval Augmentation Based on Graph Neural Network

Xi Wang, Chuantao Yin, Yuexian Li, Si Wu, Lan Wang
Sino-French Engineer School, Beihang University, Beijing, P. R. China;
Orange R&D Beijing Company Limited, Beijing, P. R. China

Abstract: Avatars are created by developers for IoT platforms by the definition of ontology in the Internet of Things (IoT) domain. It is essential to design an efficient avatar retrieval system because an IoT platform usually has millions of avatars, and developers need to use the information of avatars frequently. The conventional retrieval systems usually use the semantic similarity of queries and documents for ranking. However, it is insufficient because the relevant but not semantically similar avatars cannot be retrieved. Retrieval based on a knowledge graph can mitigate this problem but only retrieve the relevant avatars that already have links, not for unseen relations. In order to mine the possible relations between avatars and provide users with more divergent choices, we propose an avatar retrieval framework based on graph neural networks. First, we search the semantically similar avatars by the semantic similarity of query and avatar\'s name. Next, we search for those closer to the semantically similar avatars in the avatar graph to find the relevant avatars. To calculate the distance of the avatars in the graph, we use graph neural networks to get a node embedding for each avatar and calculate their distance by the cosine similarity of their node embeddings. Finally, we integrate the semantically similar avatars and the closer avatars in the graph as the final results. We conduct experiments on large datasets, which prove the effectiveness of our proposed method.

Accessible Education: Using Augmented Reality for Hearing-Impaired Students

Uday Upreti, Wazahat Attar, Mostafa Radwan, Fay Zhong
California State University, East Bay, California, USA

Abstract: The use of Augmented Reality (AR) to achieve educational inclusion has not been deeply explored. The systematic review describes the current state of using AR as an educational technology that takes into consideration the needs of all students, including those with disabilities. It is done through the analysis of various factors, such as AR's advantages, its practice, challenges, limitations, and scopes in the educational field. It further reviews the impact of using AR in learning scenarios that involve students with diverse educational needs. In this paper, we discuss the possibility of using AR technology to implement a system that interactively transcribes audio to text in real-time to assist diverse learners, especially those hearing-impaired students.

Multi-Agent Approach Promoting Interoperability between Internet of Things Applications

Patrick Pierre, Roger Nkambou, Pierre Poirier
Department of Computer Science, University of Quebec in Montreal (UQAM), Montreal, Canada

Abstract: The Internet of Things (IoT) is a promising paradigm that allows the integration of several technologies and communication solutions. According to the literature consulted, the Internet of Things (IoT) represents the next evolution of the Internet and makes it possible to considerably improve the ability to gather, analyze and render data. Already, heterogeneous and distributed services are set up by companies that will use different types of applications, all areas combined. However, the IoT raises many challenges related to the high number of objects to be considered, its high dynamicity, as well as the high heterogeneity of data and communication systems involved. Most researchers have proposed the creation of a unified standard for the integration, storage, and processing of data collected from different sources. Besides, the future expenditure in the IoT is at the level of integration and interoperability of application solutions. Therefore, it is crucial to put in place new systems and techniques in order to develop applications capable of meeting the technological requirements faced by the IoT. Interoperability between these applications remains one of the most relevant IoT issues, which needs to be addressed in detail. In this paper, we proposed and developed a MISMA-IoT (middleware via multi-agent systems for IoT application) middleware solution based on the technology multi-agent systems that could help solve interoperability issues between Internet of Things applications. In other words, our proposed middleware allows resolving the problems of interoperability and communication at the level of IoT applications in a smart city. Keywords-SMA; IdO; Smart city; Artificial intelligence.

OpenLabel: An Open-Source Media Labeling Web Browser Extension

Jason VanTol, Maxim Nikitin, Noah Helzerman, Mark Allison, Matthew Spradling
Computer Science and Information Systems, University of Michigan-Flint, Michigan, USA

Abstract: Fake news is a prevalent issue online, with far reaching consequences in society. Research has been done to classify fake news, detect it online, and provide users with labels warning them of it. This paper presents an open-source solution to displaying labels, extracting the content of a web page, and populating the label with useful information gathered from it. OpenLabel is a free and open-source media labeling Chrome extension available at: http://openlabel.xyz. The labels generated provide scored categories to inform the user of different factors indicative of fake news. This paper details how researchers and end users can make use of the software, the external packages used, and core design choices for the software architecture.

Delay-based IoT Clustering Method for Massive IoT Systems

Yunseok Chang, Seo J. Chang, Boyeon Kim
Department of Computer Engineering, Daejin University, Republic of Korea;
Department of Computer Science, University of California, Irvine, California;
AI Academy, ICT Innovation Square, Republic of Korea

Abstract: In an environment where multiple IoT nodes are clustered in groups, the dynamic clustering method can reconstruct IoT nodes into logical groups as needed, but collective control of logical groups of Massive IoT, where equipment is installed in a dense form, does not guarantee optimal control performance. In a cluster, group-level IoT control is delayed in the delivery or completion of group control commands according to grouping, but when clustering is performed on a logical group basis, IoT nodes are grouped on a user side rather than on a physical connection between IoT nodes. Therefore, when IoT nodes included in the cluster are configured as logical groups to perform dynamic clustering, each logical group needs to be re-divided into subgroups so that collective control commands can be delivered and executed within a set delay time. In this work, we present an optimal dynamic clustering method that can minimize group control latency in logical groups. To this end, this work presents a Delay-Based IoT Clustering algorithm based on DBSCAN algorithm, one of the unsupervised learning methods, when dividing clusters into logical groups according to user needs. In the DBIoTC, the communication delay and spatial distance were taken by each node in a logical group to deliver a group control command from a core node of a subgroup are expressed as two-dimensional distance values, and a group control command is configured in parallel using a DBSCAN algorithm. Multiple subgroups can be clustered according to parameters given in the algorithm, and a simple simulation experiment shows that the number of cases in which a subgroup is clustered from one logical group and the communication delay for passing group control commands for each case depends on the number of minPts values. Therefore, based on the results of this study, for dynamic clustering-based logical groups, the parameter value of the DBSCAN algorithm can be appropriately set to expand to a study on how to ensure optimal massive IoT control delay.

Analysis of a Formal Model for Social Groups

Emanuele Covino, Giovanni Pani
Dipartimento di Informatica, Universita di Bari, Italy

Abstract: We summarize the fundamental ideas of our ongoing interdisciplinary project on the definition and the analysis of social networks' models and properties; in particular, we suggest that some protocols for mobile networks that are largely adopted in computer science could be used to provide formal definitions of social structures and to analyze the related properties. We introduce a tool for the simulation and evaluation of one of these protocols.

Creating a Virtual Reality Based Shopping System

Khadija Hasta, Harsh Jobanputrua, Pravesh Jain, Fay Zhong
California State University East Bay, Hayward, California, USA

Abstract: One of the most critical features of electronic commerce (e-commerce) is that it allows individuals to purchase a wide range of products without leaving home. However, many people visit e-commerce websites but never make a purchase. One of the most fundamental problems is that consumers lose the urge to shop when faced with an old e-commerce marketplace consisting of a list of good catalogs. This paper presents a method for envisioning a Virtual Reality (VR) shopping system, explores the key factors for building a VR shopping system, and describes the current state of network-enabled technologies. When we combine the 3D vibe of a typical shopping mall with the qualities of e-commerce, it sparks shopping interest and gives e-commerce a new life. It is undeniable that the application of VR in e-commerce has a bright future.

Performance Improvement Scheme of Hyperledger Fabric Blockchain Using Parallel Preprocessing

Ho-Se Moon, Jae Geun Song, Ju Won Kim, Cheol Hun Park, Ju Wook Jang
Department of Electronic Engineering, Sogang University, Seoul, Republic of Korea

Abstract: The infinite loop problem can be a fatal problem in some programming languages. Hyperledger fabric solves the infinite loop problem with chaincode timeout. However, there is still a problem that the endorsing peer must execute chaincode having infinite loop problem. Multiple DApps can send requests to an endors-ing peer. The endorsing peers may not be able to process incoming chaincodes quickly because they have to execute chaincode having infinite loop problem. This paper proposes a method to improve the perfor-mance of the network by improving the problem that chaincodes with infinite loop problems should be executed by endorsing peers through preprocessing.

Visualization of Precise 3D Modelling of City-scale Scenes in Virtual Reality Using Photogrammetry

Krish Vora, Aditya Badhe, Tushar Pawar, Fay Zhong
California State University East Bay, Hayward, California, USA

Abstract: Photogrammetric models are critical for enabling digital versions of cities that can accurately display geographic information when loading and visualizing these city-scale scenes. However, existing methods require excessive computation and memory usage when the geographic area viewed changes rapidly in real-time rendering. In this paper, we discuss data-based 3D rendering and visualization using the Unreal Engine for Virtual Reality (VR) environment. Unreal Engine is one of the most popular 3D software in the industry today, featuring high-fidelity real-time environments and a complete set of tools for photorealistic visuals and real-time immersive experiences. Developing digital replicas of cities requires massive datasets and photogrammetry to accurately model 3D objects. In this study, we present a method that fuses multi-source geospatial data blending to enable immersive and interactive experiences in VR using photogrammetry.

Developing Augmented Reality Based Learning Tools for Organic Chemistry

Sankalp Kasale, Ashriel Waghmare, Tapasya Solu, Fay Zhong
California State University East Bay, Hayward, California, USA

Abstract: The comprehension of the microscopic world has always been the focus and challenge of Chemistry learning. The imaginative abilities of most high school students are still immature. As a result, they were unable to properly visualize the microstructure at the beginning stages of Chemistry learning. This research addresses the topic of "Composition of Substances" in high school Chemistry curriculums and involves the design and implementation of Augmented Reality (AR) based learning tools for Organic Chemistry. The proposed AR application will allow students to interact with 3D models of atoms through AR headsets, moving atoms together or breaking them apart to form new compounds.

Monitoring Integrity of IoT Network Topography using a Blockchain Network

Naresh Adhikari, Mahalingam Ramkumar
Department of Computer Science, Slippery Rock University of Pennsylvania, Slippery Rock, Pennsylvania, USA;
Department of Computer Science, Mississippi State University, Mississippi, USA

Abstract: The topographical integrity of an Internet of Things (IoT) network is a property that guarantees that (non)-malicious entities do not (un)willfully alter the 2D/3D spatial layout of network nodes. This property enhances the usefulness of the sensor data from an IoT network. A novel protocol to maintain and determine the topographical integrity of an IoT network is proposed to guarantee that sensors in an IoT network are producing valuable data from the spatial locations where they are supposed to be.

Performance Analysis of Hyperledger Fabric Version 2.2

Cheol Hun Park, Ho-Se Moon, Ju Won Kim, Jae Geun Song, Ju Wook Jang
Department of Electronic engineering, Sogang University, Seoul, Republic of Korea

Abstract: Blockchain applications can improve fault tolerance, traceability, and transparency compared to central-ized solutions. However, the industry is concerned that the current blockchain will not meet the needs of the company, such as throughput or latency. Hyperledger Fabric is a modular architecture platform for de-veloping blockchain solutions and applications, one of Hyperledger's projects. Hyperledger Fabric was first introduced as v0.6, released v1.4 in January 2019, and v2.2 in July 2020. This article compares and ana-lyzes the performance of the two versions of Hyperledger Fabric v1.4 and v2.2. The performance of both versions is evaluated in terms of latency and throughput while changing the number of transactions. As a result, Hyperledger Fabric v2.2 was found to exceed v1.4 in several areas.

Low Heat Multi-Core Assignment Scheme by Considering CPU Overloading in Big.LITTLE Architectures

Min-Ho Han, Young-Bae Ko, Sung-Hwa Lim
Department of AI Convergence Network, Ajou University, Suwon, South of Korea;
Department of Software, Ajou University, Suwon, South of Korea;
Department of Multimedia, Namseoul University, Cheonan, South of Korea

Abstract: ARM has released a big.LITTLE architecture that enables efficient use of the heterogeneous multi-core. Higher heat generated by the CPU may incur higher energy dissipation. Recently, high-performance applications such as AR and 3D have caused smartphone overheating problems. This paper proposes a multi-core assignment scheme for reducing CPU heat generation while guaranteeing the required performance. We shows that CPU utilization and heat generation of off-the-shelf smartphones are closely related by conducting experimental evaluations The proposed scheduling confirmed a 15% lower heat generation than the basic method.

Blockchain-Based IoT-Enabled Fish Supply Chain Framework

Shereen Ismail, Hassan Reza
School of Electrical Engineering & Computer Science, University of North Dakota, Grand Forks, USA

Abstract: Technologies such as blockchain and Internet of Things (IoT) revolutionize supply chain systems by incorporating immutability, transparency, and traceability mechanisms that promote secure data sharing and interactions among various stakeholders in trustless environments. To this end, we propose a blockchain IoT-enabled fish supply chain framework to securely ensure efficient tractability and traceability of fish products through supply chain. More precisely, the proposed framework should be able to provide valuable and yet timely information that can be used to track and trace the fish product throughout harvesting, processing, packaging, shipping, distribution to the final delivery. Such system adds authenticity of the product and promote anti-fraud capability in supply chain systems with the integration of machine learning.

The 21st International Conference on Wireless Networks (ICWN'22: July 25-28, 2022, USA) https://american-cse.org/csce2022/conferences-ICWN https://american-cse.org/csce2022/

A Hybrid Algorithm to Enhance Wireless Sensor Networks Security on the IoT

Ntebatseng Mahlake, Topside E. Mathonsi, Tonderai Muchenje, Deon Du Plessis
Tshwane University of Technology, Pretoria, South Africa

Abstract: The Internet of Things (IoT) is a futuristic technology that promises to connect tons of devices via the internet. As more individuals connect to the internet, it is believed that communication will generate mountains of data. IoT is currently leveraging Wireless Sensor Networks (WSNs) to collect, monitor, and transmit data and sensitive data across wireless networks using sensor nodes. WSNs encounter a variety of threats posed by attackers, including unauthorized access and data security. Especially in the context of the Internet of Things, where small embedded devices with limited computational capabilities, such as sensor nodes, are expected to connect to a larger network. As a result, WSNs are vulnerable to a variety of attacks. Furthermore, implementing security is time-consuming and selective, as traditional security algorithms degrade network performance due to their computational complexity and inherent delays. This paper describes an encryption algorithm that combines the Secure IoT (SIT) algorithm with the Security Protocols for Sensor Networks (SPINS) security protocol to create the Lightweight Security Algorithm (LSA), which addresses data security concerns while reducing power consumption in WSNs without sacrificing performance.

An Efficient Transmission Method of WSN in the Common Duct

Jongkyu Choi, Ki Young Lee
Department of Information and Telecommunications Engineering,
Incheon National University, Incheon, Republic of Korea

Abstract: The scope of this study is to use the functional play equipment as a medium which can improve depression and senile dementia that may be caused by the decline of the cognitive ability and strengthen ties between generations, and to obtain empirical results data on cognitive function improvement especially by “digital function convergence that can stimulate four senses in an analog-based board game”.

Selection Criteria for Cost-Effective Network Technologies for Rapid Expansion of Telecommunication Services in Rural Areas of South Africa

Marcel Ohanga Odhiambo, Weston Mwashita
Mangosuthu University of Technology, Durban, South Africa;
Vaal University of Technology, Pretoria, South Africa

Abstract: Rural South Africa has always been underserved in terms of telecommunications services. For numerous reasons, this condition has persisted into the era of broadband connectivity. Market failure, the slow pace of implementing SA Connect (the national broadband policy adopted by Cabinet in 2013), the splitting of the old Department of Communications, and a period of political leadership changes may all be blamed for the problem, despite the National Development Plan's vision of a dynamic and connected vibrant information society and a knowledge economy that is more inclusive, equitable, and prosperous. Poverty, inequality, and unemployment continue to be a harsh reality for the rural poor. To address all aspects of the digital ecosystem, not only access, decisive and holistic initiatives are required. The scenario necessitates stakeholders rethinking how the digital divide might be bridged. From a policy and regulatory standpoint, it necessitates novel approaches and creative thinking.

Hybrid Modulation Scheme for Reducing Peak to Average Power Ratio in 5G Based on Spatial CE-OFDMA Signal

Bertille Auyane Ngouessy ep Ovono, Temitayo Otunniyi, Marcel O. Odhiambo
Faculty of Engineering, Vaal University of Technology, Gauteng, South Africa;
Faculty of Engineering, Mangosuthu University of Technology, Durban, South Africa

Abstract: The alarming growing rates of data traffic in mobile communication systems demand the high utilization of high spectrum bands. This expected tremendous increase of mobile traffic from 5G technology has an impact on power consumption. Because of the problem of energy consumption in 5G many multicarrier modulation schemes are being studied as a part of the 5G physical layer solution to meet the requirement of 5G which aims for low latency and complexity. How-ever, those waveforms are limited by Higher Peak to Average Power Ratio (PAPR). To reduce this PAPR, various approaches have been studied based on the Constant envelope method. In this paper, a hybrid modulation scheme called Spatial Modulation Orthogonal Frequency Division Multiplexing Access (SM-CE-OFDMA) based on Constant-Envelope is used which significantly improves the performance of the system under AWGN and also reduces PAPR by 3.01dB. The design was simulated on the MATLAB programming. A comparative study is conducted to obtain the BER and SNR of each module transmitted scheme.

IPv4/IPv6 Multifaceted-based Comprehensive and Comparative Study

Khadeejah Hossain, Stefanie Markovski, Nick Perkins, Michael DePalma, Spencer Bucknam,
Westly Durkee, Olyana Mushe, Mohammed Mahmoud
Department of Computer Science and Engineering, Oakland University, Rochester, Michigan, USA;
Department of Computer Science & Information Systems, Minnesota State University, Moorhead, Minnesota, USA

Abstract: Internet Protocol version 6 (IPv6) is the most recent version of the Internet protocol which allows data transfer over the Internet. IPv6 is an improvement over IPv4. Most notably, IPv6 has a larger address size consisting of 128 bits, in comparison to IPv4 which contains 32 bits. We are quickly running out of addresses in the IPv4 namespace, however, there are vastly more addresses available in IPv6 which resolves this issue. The push for IPv6 is fueled by the limitations of IPv4 and the need for more robust and secure addresses. In this paper, we will explore the different characteristics of IPv4 and IPv6, such as security, performance/quality of service, portability, structure, IPv4 to IPv6 transition, and the applications of IPv6 in industry. Many of the vulnerabilities from IPv4 are carried into IPv6, and IPv6 differs from IPv4 in that it has more address security. Although IPv6 corrects the critical issues with IPv4, it has adopted many of the vulnerabilities that do not relate to those issues, as well as creating new ones.

Minimal Sleep Delay Driven Aggregation Tree Construction in IoT Sensor Networks

Van-Vi Vo, Duc-Tai Le, Hyunseung Choo
Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, South Korea;
College of Computing and Informatics, Sungkyunkwan University, Suwon, South Korea

Abstract: Data aggregation is a fundamental technique in wireless sensor networks (WSNs) in which sensory data collected by intermediate nodes is merged by in-network computation using maximum, average, or sum functions. Because sensors run on batteries, energy conservation is a critical issue. Duty cycle is a well-known energy-saving mechanism in WSNs, but it causes data aggregation latency to increase. As a result, the use of multichannel technology allows more sensor nodes to send data simultaneously, reducing data aggregation latency. We investigate the minimum latency aggregation scheduling problem in multi-channel duty-cycled IoT sensor networks in this paper. We propose a scheduling scheme that first constructs an aggregation tree based on sensor node sleep delay, then improves parallel transmissions by scheduling all eligible nodes in the constructed aggregation tree to enhance data aggregation. Based on extensive simulation experiments, our proposed approach lowers the aggregation delay by at most 61% compared to a novel approach.

Calculation of Existence Probability of a Tracking Target for Agent-based Human Tracking

Kozo Tanigawa, Masaru Shiozuka, Kenichi Takahashi
Development Department B, Mitsubishi Electric Software Corporation, Kube, Japan;
Graduate School of Sustainability Science, Cross-informatics Research Center, Tottori University, Tottori, Japan

Abstract: Surveillance systems for security are widely used in various facilities, such as in companies and schools, to prevent crime. In such systems, operators monitor the information from sensors, such as cameras and beacons. Therefore, we propose an agent-based human-tracking system. However, in a system in which the sensor is a camera, the target cannot be accurately determined. This can cause detection and tracking mistakes. Therefore, we propose calculating the existence probability of a target. The existence probability can be calculated by combining the traveling route of the targets with images extracted by image processing. The system can track a target more precisely and effectively based on the existence probability. Simulation results indicate that the calculation of the existence probability can achieve effective and accurate tracking.

DDoS Attacks and Botnet Detection Mechanisms Sharing Cross-planes Information in the Network: A Conceptual Security Framework

H.D.N.S Priyankara, Sureswaran Ramadass
Institute of Postgraduate Studies, Malaysia University of Science & Technology, Selangor, Malaysia

Abstract: Distributed Denial of Service (DDoS) attacks are among the most serious security issues faced by security experts. Most of the recently developed DDoS at-tack detection systems are based on Software Defined Network (SDN) controllers, which lack networkwide information-sharing capabilities and suffer from computational overheads, and detection delays. This paper proposed a novel conceptual framework for SDN-based DDoS attacks and detection of Botnet from the IP spoofed information shared across cross-planes in the network. The proposed framework investigates the traffic flow in the network via Traffic investigating agents (T-agent) and DDoS sensors installed in the network. This paper presents a novel DDoS sensor architecture and algorithm that can be used to identify IP spoofs at the source-end . It also presents a novel architecture of the Botnet identifying mechanism and its algorithm using cross-plane information.

Mobile Botnets on Wireless Network

H.D.N.S Priyankara, Sureswaran Ramadass
Institute of Postgraduate Studies, Malaysia University of Science & Technology (MUST), Selangor, Malaysia

Abstract: Mobile Botnet is an interconnected collection of smart devices involved in cyber-attacks. It is capable enough to launch various types of cyberattacks, including fraud, phishing, stealing bank credentials, and distributed denial-of-service (DDoS). Although there are surveys on the Botnet, there is no comprehensive study on mobile Botnet architectures. This paper presents fourteen architectures of mobile Botnets in the mobile network. The paper compares existing mobile Botnets covering Mobile Botnet Architecture and Protocols used to implement Command and Control servers (C&C), the architecture of the Command-and-Control Server, the control style of the C&C, characteristics, and detectability. Finally, presents the evolution of the Botnet architecture and evaluation of each architecture and resilience against the counter-defense mechanism.

Extended RTS/CTS Control Protocol for More Concurrent Data Message Transmissions in Wireless Ad-Hoc Networks

Yudai Yamazaki, Hiroaki Higaki
Department of Robotics and Mechatronics, Tokyo Denki University, Tokyo, Japan

Abstract: RTS/CTS control has been introduced to wireless LAN protocols supporting wireless ad-hoc networks. It enables collsion avoidance even data messages transmitted even by multiple hidden wireless nodes and improves performance due to less re-transmissions of data messages. However, a conventional RTS/CTS control protocol is so strict that concurrent transmissions of data messages are too restricted. This is because the protocol is designed based on a simple wireless signal transmission model, i.e., a disk model. Based on a more complex however realistic wireless signal transmission model in consideration of attenuation of wireless signal power, it is possible to design a more complex however performance-improving protocol that allows concurrent data message transmissions among neighbor wireless nodes. This paper proposes an extended RTS/CTS control protocol in consideration of the capture effects. It provides weaker restrictions on concurrent data message transmissions without additional collisions of control and data messages. Here, two different level NAVs (Network Allocation Vectors) soft and hard NAVs are introduced and they enables autonomous and distributed control for concurrent data message transmissions in compatible with collision avoidance.

Utilizing Explainable AI for Lightweight Ranging on UWB-enabled Mobile Devices

Ji-Woong Park, Young-Bae Ko
Department of AI Convergence Network, Ajou University, Suwon, South Korea

Abstract: Traditional Ultrawideband (UWB) ranging schemes show significant ranging error in Non-Line-Of-Sight (NLOS) conditions. Recent studies introduce deep learning-based approaches to mitigate such NLOS error. However, a deep learning approach is difficult to operate in real-time on mobile devices due to limited computing resources. To solve this problem, we propose a lightweight and accurate ultrawideband ranging scheme based on Explained Artificial Intelligence (XAI). We utilize the XAI algorithm to select important peaks from ultrawideband Channel Impulse Response (CIR). A novel lightweight classifier model is trained using the selected important peaks. In the performance evaluation, we reduce the model size by 1100x, sacrificing only 4 % accuracy

The 21st International Conference on Information & Knowledge Engineering (IKE'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-IKE
https://american-cse.org/csce2022/

Investigating Collaboration for Software Reverse Engineering

Allison Wong, Wayne Henry
Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright-Patterson AFB, USA

Abstract: Reverse engineering (RE) is a process of exploration and analysis to support software design recovery and exploit development. The process is often conducted in teams to divide the workload and take full advantage of engineers' individual expertise and strengths. Collaboration in RE requires versatile tools that can match the environment's unpredictable and fluid nature. While studies on collaborative software development have indicated common best practices and implementations, similar standards have not been explored in reverse engineering. This research conducts semi-structured interviews with reverse engineering experts to understand their needs and solutions while working in a team. The results describe an array of major challenges that are addressed by employing tools such as issue tracking software, shared workspaces, and version control systems. Such tools support documentation and continuity, while mitigating redundancies in concurrent work. Though the value of these tools is acknowledged by the experts, seamless workflow integration remains a challenge. The identification of current needs and practices offers additional opportunities for collaborative tool developers to aid reverse engineers.

Ensemble Models for Classification

John Tsiligaridis, Manuel Anaya
Math and Computer Science Department, Heritage University, Toppenish, Washington, USA

Abstract: An ensemble method is viewed as a compound model. The purpose of such a model is to achieve better predictive performance. The attempt is to tune predictions to observations by decreasing model variance, and bias. Stacking is an ensemble method, involving the base models and the metamodel that combines their predictions. The need of increasing accuracy can be covered by the combination of models to form a single prediction. To this end, firstly an ensemble model of a Random Forest (PARF), based on the Projection Algorithm (PA), is developed. Secondly, two more algorithms WG and WO are presented for the weighted ensemble method using stalking. Promising results based on accuracy of the proposed models are obtained.

System Dynamics Applications in Emergency Department (ED) for Waiting Time Improvement for Medical Care

Stephane Copin, Selmin Nurcan, Tamim Hammadi
University of Paris 1 Pantheon, Sorbonne Paris, France

Abstract: Emergency department (ED) overcrowding and long wait times as a result of continued growth in demand and non-emergency visits represent a significant proportion of visits to most ED. The use of short circuits with the addition of a reception and guidance doctor (RGD) at thes triage seems to be interesting way to reduce the pressure on emergencies and reduce the average waiting time for the medical care of patients. The objective is to introduce systems thinking and modeling of system dynamics in the ED of the Greater Hospital of East Francilien (GHEF) in France as techniques to assess the impact of the implementation of a RGD on ED patient wait times and explore ways to improve patient flow and system capacity. This article provides a detailed overview of a model of systems dynamics that focuses on the ED; it discusses the strengths and weaknesses we found in its application and suggests some opportunities for its development and use in the future. The simulation results show that the establishment of a short circuit made it possible to significantly reduce the waiting time for medical treatment by 18% on average.

X-PoSuRe: Pollution Super Resolution for Finegrained Inference using Multimodal Data Fusion and Deep Learning

Srinivas Devarakonda, Senthil Chittaranjan, Badri Nath
Rutgers University, New Brunswick, New Jersey, USA;
Anna University, Chennai, India

Abstract: This study proposes X-PoSuRe - a neural network based regression model for pollution super resolution that is trained to infer finegrained pollution information from coarsegrained pollution measurements akin to image super resolution, where a generative model creates high resolution images from low resolution images. The X-PoSuRe model uses Nitrogen Dioxide (NO2) as the pollutant of choice and uses other covariates like meteorological data, traffic data, construction activity, accident information, large scale fire incidents, and building footprint information for pollution super resolution. The model is built using an ensemble of neural network models and trained using gradient boosting technique. The results show that finegrained NO2 concentrations can be inferred from the covariates with high accuracy using this model. The proposed X-PoSuRe model provides a promising new and novel method for pollution super resolution from existing low resolution data sources without the need for deploying expensive measurement equipment over a large area.

A Conceptual Solution Approach for Artificial Intelligence Adaption along the Student Lifecycle in Higher Education Institutions

Christian-Andreas Schumann, Nico Kling, Frank Otto, Claudia Tittmann
Institute for Management and Information, West Saxon University Zwickau, Zwickau, Germany

Abstract: In this paper, the authors describe results of an ongoing project in the field of artificial intelligence (AI) in higher education (HE) at the West Saxon University Zwickau (WHZ). The impetus for AI in HE is presented, followed by a description of the environment, specifically problems and a solution approach, and the knowledge base is shown with the student lifecycle management of the WHZ. Results of the environment and the knowledge base are consequently used to build a theoretical implementation of a smart conversational agent, as well as a concept about linking distributed information systems within a university by using an ontology.

Traffic Light Control Using Fuzzy Logic

Rudy Argueta, Brian Duenas, Andy Osorio, Giovanni Sandoval Ochoa, Ruting Jia
Electrical and Computer Engineering, California State University Northridge, Northridge, CA, USA

Abstract: This paper presents the use of Fuzzy Logic to help increase the efficiency of Traffic Lights. Fuzzy Logic originally introduced by Dr. Lotfi Zadeh an electrical engineer, mathematician and scientist. The Fuzzy Logic Algorithm is created in order to create a new timing mechanism. This timing is applied to a busy intersection and then compared to the Department of Transportation's(DOT) timing manual. The design of the Fuzzy Logic Traffic Control system is tested using the Fuzzy Logic Toolbox and AnyLogic Software. The use of these simulations allowed for a new implementation of traffic light delay values, as opposed to what is already in place. AnyLogic also utilized to simulate traffic and collect data using the DOT timing manual and the new Fuzzy Logic Timing. Data from both the DOT and the Fuzzy Logic simulations are compared to one another. Six cases are examined in this study and each cases varied using the traffic density and speed variables. The fuzzy logic algorithm on average outperformed the DOT by approximately 4.93 seconds for the 6 cases examined.

Of Polygons, Curtains, and Corridors in Multidimensional Grid Spaces

Peter Baumann, Dimitar Misev, Bang Pham Huu
Jacobs University, Bremen, Germany

Abstract: Multi-dimensional arrays form a basic category in programming, and incidentally they also appear frequently in Big Data - satellite image timeseries, climate analysis, human brain data, and cosmological simulations are just a few out of many use cases. In business data, OLAP datacubes are common since long while in science and engineering they have become popular only recently, mostly with Earth data. Among the specific new requirements in these domains is support for polygon clipping on massive multi-dimensional raster data. In this paper we present a novel operator for polygon/raster clipping in Array Databases. It not only accomplishes classical clipping, but additionally allows simplified extraction of non-planar regions ("curtains") and extraction of data along an arbitrary path ("corridors"). The algorithms are streamlined to the processing of partitioned ("tiled") arrays. We present the concepts and the array query operator together with extensive benchmarks considering lines from 2-D and 3-D, polygons from 2-D, codimension 1 affine subspaces from 3-D, and curtains from 3-D data with both linear and polygonal cross sections.

Scenario-based Perspectives of Popular Ethical Theories on Cybersecurity

Natnael Birhanu Teferi, Nahom Atsbha Gebreyohannes, Megan Steinkuehler,
Maya Kruger, Nathan Ferrell, Nadin Qaqos, Mason Voxland, Mohammed Mahmoud
Department of Computer Science & Information Systems, Minnesota State University, Moorhead, Minnesota, USA

Abstract: Ethics within the field of cybersecurity is a topic that is often debated. There have been many attempts to standardize ethics within cybersecurity, but in our opinion it is either impossible to do so, or has yet to be done. In this paper, we will attempt to contribute to this discussion by considering five scenarios, looking at each one through the lenses of Kantianism, Act/Rule Utilitarianism and Social Contract Theory. These examples will provide a realistic scenario so that a specific situation is being considered instead of simple theories, like in real life. This will show that ethics and morality, especially within cybersecurity, are not so cut and dry, and that much more discussion needs to take place before a single viewpoint can be implemented.

The Ethical Risks and Challenges in Big Data

Mohammed Mahmoud, Dafer Alali, Surah Aldakhl
Department of Computer Science and Engineering, Oakland University, Rochester, Michigan, USA;
Department of Electrical and Computer Engineering, Oakland University, Rochester, Michigan, USA;
Department of Computer Science and Information Systems, Minnesota State University, Moorhead, Minnesota, USA

Abstract: Big Data refers to the digital sensation that authorizes the collection and use of an enormous amount of data derived from both man and machine. All these datasets are characterized into six categories: high volume, velocity, variety, veracity, variability and complexity. Big Data allows different businesses to collect, analyze, and use an individual's information at a rapid pace. However, it also allows for the power of invading someone's privacy. In either situation, Big Data creates the discussion of ethical issues correlated to the sharing and usage of data. In this paper, I will discuss the opportunities, risks and challenges of ethics in Big Data, as well as four different ethical theories, which include Kantianism, Act/Rule Utilitarianism, Social Contract Theory and Virtue Theory. These theories help frame our mindset on ethical issues.

Comparative Analysis of Linguistic Processors based on Machine Learning to Predict Suicidal Sentiments on Twitter

Lizethe Guadalupe Reyna Moran, Francisco Javier Luna Rosas, Julio Cesar Martinez Romo, Ivan Castillo Zuniga, Marco Antonio Hernandez Vargas, Ricardo Mendoza Gonzalez
Computer Science Department, Inst. Tec. Aguascalientes, Mexico;
Computer and Systems Department, Inst. Tec. Aguascalientes, Mexico;
Electrical and Electronic Engineering Department, Inst. Tec. Aguascalientes, Mexico

Abstract: For health professionals, suicide is a complex issue, in addition it has been shown that many deaths by suicide are preventable. The present work aims to conduct a comparative analysis of four types of linguistic processors (LIWC, TextBlob, Vader and NRC) based on evaluating the accuracy of machine learning models with characteristics of the linguistic processors. This to classify texts related to suicide on the social network Twitter. The accuracy of the machine learning models using features given by the linguistic processors was on average exceptionally good, mostly approaching 100%. The highest accuracy score was using the TextBlob linguistic processor with 99.86%.

Promoting the Flow of Knowledge in Project-based Academic Organizations based on the Zachman Framework

Azam Ranjbar, Peyman Akhavan, Raheleh Khatami, Lila Rajabion, Reza Hosnavi
Malek Ashtar University of Technology, Tehran, Iran;
Qom University of Technology, Qom, Iran;
Iran Knowledge Management Association, Tehran, Iran;
State University of New York, New York, USA

Abstract: Today, identifying and improving the flow and proper management of knowledge in an academic organization is one of the principal means of creating a competitive advantage for that organization, especially in project-based organizations where a large amount of knowledge is circulating. Given the importance of knowledge flow in project-based organizations, this paper will attempt to provide a model for improving knowledge flow using the Zachman modeling framework. This model is designed using the conceptual framework of knowledge flow modeling. The framework provides a matrix whose rows include areas of scope, organization, command and control, role, and technology. The columns of the framework are the aspects of knowledge flow modeling that explain the purpose of knowledge flow, transferred knowledge content, factors involved in knowledge flow, types of vertical and horizontal knowledge flows, time of knowledge flow occurrence, and mechanism of knowledge flow in the model. The proposed model will be able to improve the flow of knowledge in project-based research centers.

AAGCGN: An Attention-Based Adaptive Graph Convolutional Gating Network for Traffic Prediction

Renyi Chen, Huaxiong Yao
Central China Normal University, Hubei, P. R. China

Abstract: Predicting road traffic conditions is of great significance to the development of society, but road conditions have complex spatial and temporal correlations, so the modeling of traffic prediction has great challenges. In this paper, we propose An Attention-based Adaptive Graph Convolutional Gating Network (AAGCGN) to overcome the aforementioned challenges. The model adopts the encoder and decoder architecture of attention mechanism to improve the parallel computing capability. A residual network is used in the encoder and decoder to stack multi-layer of spatial-temporal blocks, which include adaptive graph convolution and temporal attention to extract complex spatial correlations and non-linear temporal dependencies. Between multi-layer spatial-temporal blocks, we introduce a gate function to control the flow of residual information. Extensive experiments on two real-world road network traffic datasets, METR-LA and PEMS-BAY, show that the AAGCGN achieves the state-of-the-art results.

An Empirical Study on Feature Selection for Prediction of Online Shopping Purchases

Hank Helmers, Huanjing Wang
Western Kentucky University, Bowling Green, Kentucky, USA

Abstract: Online shopping has expanded dramatically with the popularization of digital services such as Amazon. In this study, we explore the data collected during customers' online-shopping sessions to predict customers' purchases better and understand the online shopping data. Experiments are conducted on an imbalanced dataset from a real-world online shopping website. Thus, feature selection allows an intelligent selection of only the most influential features, which can improve classifier performance. We investigate five feature ranking evaluators and build classification models using three classifiers. Results demonstrate that the Information Gain feature selection and the Multilayer Perceptron classifier performed the best.

The Role of Predictive Analytics in Individual and Public Healthcare

Amitava Karmaker
University of Wisconsin-Stout, Menomonie, Wisconsin, USA

Abstract: In the era of big data applications, predictive analytics plays a vital role in understanding various issues in the world. As a matter of fact, the applicability of predictive analytics encompasses fields such as social studies to technological applications predicting diseases within the healthcare industry. Historical data and trend analysis can help identify an effective prediction process to determine the occurrences of diseases and even a pandemic. In this paper, I will highlight how the use of data in disease prediction can be adequately solved through predictive analytics in big data. The early detection of any unprecedented catastrophe will help us avoid disruptions ranging from economic disruptions to social disruptions.

Unsupervised Reference Correction Methods for a Data Washing Machine

Fumiko Kobayashi, John R. Talburt
Donaghey College of Science, Technology, Engineering, and Mathematics,
The University of Arkansas at Little Rock, Arkansas, USA

Abstract: A Data Washing Machine (DWM) is an unsupervised system accepting any type of data as an input and generates a fully cleaned, and ready to use data product. This paper describes six methods of unsupervised reference corrections that can be used in a DWM to help achieve a working system. These leverage the context of the input data and reference groupings that are determined by the DWM to make context-based corrections to string values within the references. These methods can be applied to both pre and post linked references to not only enhance equivalence decisions but also to apply Data Quality (DQ) corrections to the output data. This paper focuses on introducing the six methods and reporting the initial results observed. These results show that when the methods are applied, they provide a lift in the F-Measure for over 87% of the runs tested.

Data-Centric Content Classification of Smart City Residential Services

Ding Li, Jincheng Sun, Yan Liu, Wenjie Du
Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

Abstract: The residential services in the context of smart cities accumulate massive real-time inquiry data in natural language to describe the services in need. Such inquiry requests have diverse topics in the content and considerable variate length. Besides, the responsible departments that may handle the inquiry involve a large number of organizations, from metropolitan administration to local communities. Hence accumulated request data is primary and central to the service of accurately dispatching requests to responsible departments. The challenge is devising a data-centric approach to fit the data with SOTA models and improve the request classification accuracy. In this paper, we analyze the factors of embedding tokens, data segmentation, model structures, and classification methods. We devise a unified modelling process with multiple dataflows that combine these factors to observe their interactions. The experiment results demonstrate the compound effects and provide insights into how SOTA models respond differently to variations in these factors. The observations allow us to fine-tune the learning task at each stage and achieve a maximum of 82.4% F1-Score.

Discretization of Multi-valued Categorical Data Attributes

Ray R. Hashemi, Keith Galbraithy
Department of Computer Science, Georgia Southern University, Savannah, Georgia, USA

Abstract: There is a large class of: (1) objects' attribute named multi-valued categorical data (MVC) attributes for which each attribute value is an array of categorical data (e.g., employees' skills, products' components, properties' interior, etc.) and (2) machine learning algorithms that are applicable only on the discretized data (e.g., ID3, Rough sets, Formal Concept Analysis, etc.). Using one such algorithm to extract features from a dataset with MVC attributes demands the discretization of MVC attributes. We introduce an association-based methodology (ABM) for discretization of MVC attributes and establish the ABM validity by using a synthesized dataset and C4.5 algorithm. The results revealed the superior performance of the ABM over the D-powerset (a common approach used for such discretization).

On Searching a Approximately Sorted Array

Aditya Narasimhan, C. R. Subramanian, Sridhar Radhakrishnan
School of Computer Science, University of Oklahoma, Norman, Oklahoma, USA;
The Institute of Mathematical Sciences, HBNI, Chennai, India

Abstract: Sorting and searching with the extremely large amounts of data, these problems become challenging. In order to reduce the resources used, we introduce the idea of approximate sorting. We also define quality metrics that measure the sortedness of these approximately sorted arrays. Our goal in this paper is to develop a binary search algorithm, performed on an approximately sorted array with these metrics L given or not. We show that when L is known for maximum displacement (md) and distance (dis) metrics, the binary search can be completed with log (n/4L) + 4L + 1 and log (n/2L) + 2L + 3 comparisons, respectively, where n is the number of elements in the array. When L is unknown, the binary search can be completed with log (n/L*) + L* comparisons, where L* is the upper bound on the value of the metric.

Database Intrusion Detection and Prevention Using Query Rewrites and Execution Plans

Aspen Olmsted
Simmons University, Boston, Massachusetts, USA

Abstract: This paper discusses our implementation of an intrusion detection and prevention system to defend against database attacks. Our solution utilizes functionality that allows us to extend the database server to see a query before and after the query is parsed. Our solution can also modify the incoming request. We compare our solution to historical solutions that utilize pattern matching, randomization, and tainting techniques. We document database attack categories and develop test cases executed against our solution and implementations of historically available solutions. The database query rewrite solution has minimal overhead in the databases that support the extensibility feature. As a result, our solution scales and detects attacks better than historical solutions. We also implemented a nightly batch job to audit changes to query execution plans to detect new zero-day data-base attacks. Changes in execution plans are a good indicator of both malicious and non-malicious changes in database queries. Any execution plan changes are examined through machine learning to classify queries as malicious that the non-machine learning techniques may have missed. We tested our solution with a few open-source databases, but any database vendor could implement the methods to protect the attack surface of the database better.

Pandemic Equation and Waves of COVID-19 Evolution

Michael Shur
Rensselaer Polytechnic Institute, Troy, New York, USA;
Electronics of the Future, Inc. Vienna, Virginia, USA

Abstract: The waves of the COVID-19 pandemics caused by the new virus strains and affected by the mitigation measures, such as the introduction or removal of the quarantine, and vaccination depend on time and location. The Pandemic Equation accurately describes the pandemic metrics locally and globally by accounting for the “curve flattening” of the pandemic waves and the mitigation measures. The Pandemic Equations uses the growth constants slowing varying with time for “curve flattening” and the generalized Fermi-Dirac distribution functions for describing mitigation and anti-mitigation measures. The Pandemic Equation parameters extracted from the well-advanced pandemic curves can be used for predicting the pandemic evolution over a limited period of time and comparing different scenarios of the pandemic development. The parameter extraction for multiple locations could also allow for uncertainty quantification in predicting the pandemic evolution using the Pandemic Equation.

Multitenant Database Security with Shared Data Policy

Aspen Olmsted
Simmons University, Boston, Massachusetts, USA

Abstract: This paper discusses our implementation of a multitenant database security solution for use cases where master records are shared but related documents are secured based on transactional activity. We utilize a motivating example of the Lincoln Center for the Performing Arts in New York City. Lincoln Center has several dozen constituent organizations which have individual ownership of the data elements, but the organizations allow transactions from any partner organization. Our solution utilizes functionality that will enable us to extend the database server to see a query before and after the query is parsed. Our solution can also modify the incoming request to apply security predicates. We compare our solution to historical solutions that utilize row-level security techniques. We document the significant difference in security labels that must be maintained in traditional systems compared to our solution's lower number of labels. The database query re-write solution has minimal overhead in the databases that support the extensibility feature. As a result, our solution scales and allows better security than historical solutions. We also compare our solution to policy-based row-level security solutions. Many improvements can be implemented in the policy-based row-level security but require hardcoding the rules.

Gas Turbine Performance Enhancement Utilizing Energy-Mix for Power Generation Reliability

S. L. Braide, D. C. Idoniboyeobu
Department of Electrical Engineering Rivers State University, Nigeria

Abstract: This paper applied decomposition techniques to determine the capacity combination for fossil fuel consumption pattern. The generating capacity are decomposed into different optimization options on the view to make operational cost saving in accordance to load demand and generating capacity of 2250 MW for Afam, 2350 MW for Sapele and 3000 MW for Egbin Station. Projection plans of 20 years was used for the analysis. Five optimization was developed for cost savings plan 1, 2, 3, 4 and 5 respectively. utilizing Renewable Energy sources (solar and wind energy) with hydrocarbon sources as form of energy mix for electricity power generation are proposed in power structure, Nigeria. The consequences included changes in climatic scenarios due to burning of hydrocarbon, environmental pollution necessitate deep reduction in greenhouse gases. Since availability of Renewable Energy resources particularly wind and solar potentials of all 36 states and Federal Capital Territory are harnessed. The primary sources of information was accessed using data in selected location in each state. Determination of wind speed profile for wind-turbine generation required estimation of wind potential using Weilbull probability distribution. Similarly, solar irradiation values across several places in Nigeria of the sun-radiation potential are simulated with Hybrid Optimization Model Renewable Energy Sources (HOMER). The increasing demand for electricity supply and its availability is indispensable to manage economically following the rapid growth of population to energy consumption capacity is a major concern making the electricity sector to experienced significant implication for power system plant operation to provide basic energy services to the consumers. To achieve this goal the study will consider the application of reliability technique in order to analyse the activities of failure times of ten identical gas-turbine blade of similar make and mode of operations subjected to the same conditions in Afam power station over a period of ten years from the results obtained gas turbine blades were in their wear-out period of reliable maintenance to give out efficient performance following to the reliability of the three parameter weilbull distribution R(t) given as 0.60 this means that turbine blade actually required reliable maintenance. While, failure rate of the turbine-blade is $\lambda(t)$:0.082577/hrs, the mean time to failure (MTTF) is 13.55hours. The results obtained through the simulation of TPC windchill quality solution software estimated the parameters which shows suitable behaviour of the system components for early response for reliable maintenance. The reliability R(t), failure rate $\lambda(t)$ and meantime to failure (MTTF) were successively computed. Conclusively, the probability that the gas turbine blades under investigation will continue to be operational in service without failures is about 70% while the mean time to failure of the gas turbine blade is about 14hours.

Representative Sampling: Deriving Accurate Data Insights from Noisy Data Distributions

Neil Palleti, Saigopal Thota, Mridul Jain, Rijul Magu
Walmart Global Tech, Cupertino, California, USA; Walmart Global Tech, Sunnyvale, California, USA; Walmart Global Tech, San Jose, California, USA; Walmart Global Tech, Atlanta, Georgia, USA

Abstract: Discovering insights from data is vital to a business' ability to understand historical trends, predict future patterns, and shape business strategy. However, real world data is often noisy and sparse, and extracting accurate insights from impure data is challenging. We present Representative Sampling, a novel statistical method to generate a high-quality subset of data from high-dimensional noisy data. Our method ensures that the subset of data is clean and representative of the true data distribution, while eliminating existing impurities. Analyzing this clean subset of data leads to significantly more accurate data insights. While our use case revolves around customer data, we detail how Representative Sampling can be used in any big data field suffering from data impurities.

The 26th International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-IPCV
https://american-cse.org/csce2022/

Meta-SeL: 3D-model ShapeNet Core Classification using Meta-Semantic Learning

Farid Ghareh Mohammadi Cheng Chen, Farzan Shenavarmasouleh, M. Hadi Amini, Beshoy Morkos, Hamid R. Arabnia
School of Computing, University of Georgia, Athens, Georgia, USA;
College of Engineering, University of Georgia, Athens, Georgia, USA;
Knight Foundation School of Computing and Information Sciences,
Florida International University (FIU), Miami, Florida, USA

Abstract: Understanding 3D point cloud models for learning purposes has become an imperative challenge for real-world identification such as autonomous driving systems. A wide variety of solutions using deep learning have been proposed for point cloud segmentation, object detection, and classification. These methods, however, often require a considerable number of model parameters and are computationally expensive. We study a semantic dimension of given 3D data points and propose an efficient method called Meta-Semantic Learning (Meta-SeL). Meta-SeL is an integrated framework that leverages two input 3D local points (input 3D models and part-segmentation labels), providing a time and cost-efficient, and precise projection model for a number of 3D recognition tasks. The results indicate that Meta-SeL yields competitive performance in comparison with other complex state-of-the-art work. Moreover, being random shuffle invariant, Meta-SeL is resilient to translation as well as jittering noise.

AIC-AB NET: A Neural Network for Image Captioning with Spatial Attention and Text Attributes

Guoyun Tu, Ying Liu, Vladimir Vlassov
Department of Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden;
Norna, Stockholm, Sweden

Abstract: Image captioning is an important field across computer vision and natural language processing. We propose and present AIC-AB Net, a novel Attribute-Information-Combined Attention-Based Network that combines spatial attention architecture and text attributes in an encoder-decoder. For caption generation, adaptive spatial attention determines which image region best represents the image and whether to attend to the visual features or the visual sentinel. Text attribute information is synchronously fed into the decoder to help image recognition and reduce uncertainty. We have tested and evaluated our AIC-AB Net on the MS COCO dataset and a new proposed Fashion dataset. The Fashion dataset is employed as a benchmark of single-object images. The results show the superior performance of the proposed model compared to the state-of-the-art baseline and ablated models on both the images from MSCOCO and our single-object images. Our AIC-AB Net outperforms the baseline adaptive attention network by 0.017 (CIDEr score) on the MS COCO dataset and 0.095 (CIDEr score) on the Fashion dataset.

A New Approach for Liver and Its Tumor Segmentation in CT Image by an Improved TransUNet with Cbam

Yaqi Hu, Xiaoniu Yang, Xiaolin Tian
Faculty of Information Technology, Macau University of Science & Technology, Macau SAR, P. R. China

Abstract: Computed Tomography (CT) can provide information of liver and its tumors for patients and doctors, and help doctors in early diagnosis, treatment planning, tumor volume measurement and monitoring clinical responses processes. Nowadays, the extensive usage of CT image in medical area has shown great importance of applying automatic medical treatment method to reduce the workload of medical workers. Due to CT image shape and contrast difference, manually segment the liver and tumors images is a time-consuming and challenging work. Therefore, several research teams have already introduced neural network such as UNet into the liver and its tumors segmentation challenge. Our research also focusses on automatic liver and its tumors segmentation using an improved TransUNet. Except the attention mechanism that TransUNet used, we add one more Cbam architecture into the model. The training and testing dataset is from the public dataset named Liver Tumor Segmentation Challenge (LiTS-2017). Compared with the first paper, we have preprocessed all images in dataset. Changing the window width and window center by cutting the HU value according to each patient sample's tumor maximum and minimum HU value. Dice score of liver and tumor segmentation in training and testing is 98%, 91%. For only segment the liver, before preprocessing, the training and testing score is 98%, 89%, after preprocessing, the score is raised to 98%, 91%, respectively, from the images of LiTS dataset.

Neural Network Based Color Pixel Classification Using Random Negative Examples

DongKeun Kim
Department of Computer Engineering, Kongju National University, Republic Korea

Abstract: In this paper, we propose a color pixel classification method in an image. It is based on a neural network (NN) using random negative examples. Positive examples are given as reference images by user which is interested in colors for classifying images. Negative examples are generated by random sampling in color space. Outliers in negative examples are eliminated by using sample means from positive examples. The experimental results showed that it is very effective to detect multi-colored pixels in an image.

Order Statistics Distributions Around Edges

Ahmed Raji
Laboratory of Images, Signals and Intelligent Systems, Paris Est Creteil University, France

Abstract: Nonlinear filters offer significant advantages over linear methods in applications involving non Gaussian random processes or when the systems acting on the signals of interest are inherently nonlinear. However, while linear filters used in numerous practical applications benefit from consistent linear filtering theory tools, nonlinear filters lack a unified and universal set of tools for analysis and design. In this article, we address some analytical aspects relative to the order statistics distributions near edges, assuming that knowledge of such behaviors can provide useful help in designing order statistics based nonlinear filters.

Multi-Modality Image Inpainting using Generative Adversarial Networks

Aref Abedjooy, Mehran Ebrahimi
Faculty of Science, Ontario Tech University, Oshawa, Ontario, Canada

Abstract: Deep learning techniques, especially Generative Adversarial Networks (GANs) have significantly improved image inpainting and image-to-image translation tasks over the past few years. To the best of our knowledge, the problem of combining the image inpainting task with the multi-modality image-to-image translation remains intact. In this paper, we propose a model to address this problem. The model will be evaluated on combined night-to-day image translation and inpainting, along with promising qualitative and quantitative results.

PlantNet: Recognition of Plant Species using Deep Convolutional Network

Sitara Afzal, Imran Ullah Khan, Jong Weon Lee
Mixed Reality and Interaction Lab, Department of Software, Sejong University, Seoul, Republic of Korea

Abstract: Plant specie identification is a significant challenge in the fields of computer vision and pattern recognition. This article presents a new approach to plant specie identification, by integrating the shape and texture characteristics. First, we introduce a custom dataset of 48 categories of different species of plants and then applied our new CNN model. The proposed approach has been thoroughly evaluated on our custom dataset. We identify the plant species with the whole plant as well as only the leaf part. Our method achieves 95.02% accuracy on this custom dataset. The recognition performance of our method is better or comparable to the prior state-of-the-art plant leaf recognition method.

Automation of Luminescence Quantitation for High-Throughput Plant Phenotyping Using Image Processing and U-Net Segmentation

Mahbubur Rahman, Afroza Haque, Dinej S. Pujara, Jose Mayorga, Hong-Gu Kang, Damian Valles
Ingram School of Engineering, Texas State University, San Marcos, Texas, USA;
School of Electrical and Computer Engineering, Purdue University, Fort Wayne, Indiana, USA;
Department of Biology, Texas State University, San Marcos, Texas, USA

Abstract: Computation of luminescence in image-based plant phenotyping is a prolonged and tiresome process. A luminescence reporter strain of Pseudomonas syringae, a bacterial pathogen, was developed to expedite this quantitation in resistance trait analysis. To further facilitate the high-throughput pipeline of the plant resistance analysis, an image processing, and a deep learning model are developed to quantify the luminescence from infected plants' images with the reporter strain captured in the dark condition. Low light images of the corresponding plant were captured to perform this computation effectively. Therefore, the image segmentation technique with tuned parameters identifies the mask of the leaves from low-light images. Every image mask associated with the corresponding luminescence image computes the mean luminescence of the targeted leaf which is linearly correlated with the bacteria population in the affected plant. Conventionally, an infected plant is assessed by the leaf disc-based manual method, which is time inefficient and prone to errors. Proposed algorithms in this paper will help compute bacterial populations faster in image-based plant phenotyping.

A Comparative Study on the Application of K-NN, ANN and RBFN to Predict Severe Thunderstorm

Himadri Bhattacharyya Chakrabarty, Sonia Bhattacharya, Ajoy Kumar Chakraborty
Bangabasi College affiliated to University of Calcutta, India;
Department of Computer Science, Panihati Mahavidyalaya, West Bengal State University, Kolkata, India;
Department of Irrigation and Water Resources, Government of West Bengal, India

Abstract: Severe Thunderstorm is the extreme weather convective feature. It causes local calamities in various ways. Proper prediction with lead time is an important factor to prevent such calamities for saving the people. Here machine learning technique is applied on weather data to get proper prediction. A comparative study has been done on application of three methods of pattern recognition namely Artificial Neural Network (ANN), K-Nearest Neighbor (K-NN), and Radial Basis Function Network (RBFN) on weather data RSRW procured from Kolkata located in North-East India. The result obtained applying Radial Basis Function Network is better among three methods yielding correct prediction of 95% for severe "squall-storm" and 94% for "no storm". Prediction can have sufficient lead time of 10- 12 hours.

Unsupervised Graph-Based Object Search and Retrieval Using Airborne LiDAR Point Clouds

Yassine Belkhouche, Bill Buckles
Department of Computer Science, Missouri State University, Springfield, Missouri, USA;
Department of Computer Science and Engineering, University of North Texas, Denton, Texas, USA

Abstract: We established a framework for object search and retrieval from airborne LiDAR data. The proposed method finds and retrieves target objects using three steps. LiDAR filtering, in which we separate object points from terrain points. In the second step, we extract individual objects by clustering object points. For each extracted object, we computed a unique signature. This signature is used in the final step to retrieve objects similar to the target object. We tested the proposed method using two real-world datasets. The proposed method effectively extracted all objects similar to the target object.

Invariant Features for Multi-Component Shapes

Mohamed B. H. Rhouma, Mohamed A. Khabou
Department of Mathematics, Qatar University, Doha, Qatar;
Department of Electrical and Computer Engineering, University of West Florida, Florida, USA

Abstract: We propose a set of rotation-, translation-, and scale-invariant features for rep-resentation and analysis of multi-component images. The features are moment-inspired and encompass information about the moment of the individual components in the image and their relative position to each other. The features are tested on a set of computer-generated images and were shown to effectively capture the nature of the images.

Modified Topological Image Preprocessing for Skin Lesion Classifications

Hong Cheng, Rebekah Leamons, Ahmad Al Shami
Department of Computer Science, Southern Arkansas University - Magnolia, Arkansas, USA

Abstract: This paper proposes a modified Topological Data Analysis model for skin images preprocessing and enhancements. The skin lesion dataset HAM10000 used with the intention of identifying the important objects in relevant regions of the images. In order to evaluate both the original dataset and the preprocessed dataset, Deep Convolutional Neural Network and Vision Transformer models were utilized to train both models. After training, the experimental results demonstrate that the images preprocessed using the Modified Topological Data Analysis consistently perform better.

Palm Vein Identification and Verification Using Deep Metric Learning

Aung Si Min Htet, Hyo Jong Lee
Division of Computer Science and Engineering, Jeonbuk National, Jeonju, Republic of Korea

Abstract: Palm vein recognition is the emerging biometrics technology for personal identification. Unlike extrinsic methods such as fingerprint, face, and palmprint recognition, palm veins are remarked as reliable and secure biometric features as it is difficult to duplicate and remain unchanged throughout one's lifetime. In recent years, many researchers have proposed for palm vein recognition using state of the art deep learning approach with promising outcome results. However, most of the existing methods are primarily focused on classification problem, which is not that much applicable for real word scenario. In this research, we experimented deep metric learning using SEResNet on CASIA multispectral palm vein dataset and Additive Angular Margin Loss is used for learning highly discriminative features for palm vein representations. Moreover, You-only-look-once (YOLO) real time object detection model is also applied to extract the responsive ROI images before training the recognition model. In our study, we have achieved 98% recognition rate for the identification using Support Vector Machine classifier after training SEResNet, and 99% True Positive Rate (TPR) at the EqualError-Rate (EER) of 0.01 for the verification and deep metric learning.

A Study on the Use of Perceptual and Robust Hashing to Detect Manipulation Of Embedded Messages in Images

Sven-Jannik Wohnert, Kai Hendrik Wohnert, Eldar Almamedov, Carsten Frank, Volker Skwarek
Digital Business Processes Research and Transfer Centre,
Hamburg University of Applied Sciences, Hamburg, Germany

Abstract: Typically, metadata of images are stored in a specific data segment of the image file. However, to securely detect changes, data can also be embedded within images. This follows the goal to invisibly and robustly embed as much information as possible to, ideally, even survive compression. This work searches for embedding principles which allow to distinguish between unintended changes by lossy image compression and malicious manipulation of the embedded message based on the change of its perceptual or robust hash. Different embedding and compression algorithms are compared. The study shows that embedding a message via integer wavelet transform and compression with Karhunen-Loeve-transform yields the best results. However, it was not possible to distinguish between manipulation and compression in all cases.

Deep Learning Convolutional Neural Networks for Brain Tumour Segmentation: Enhanced Classification Loss Function (ECLF)

Komal Grover, Abeer Alsadoon, Mahmoud Bekhit, Nada Alsallami, Ahmed Dawoud,
Muhammad Siddiqi, Razwan Mohmed Salah
School of Computing Mathematics and Engineering, Charles Sturt University, Australia;
School of Computer Data and Mathematical Sciences, Western Sydney University, Australia;
Kent Institute Australia, Sydney, Australia; Asia Pacific International College (APIC), Australia;
Computer Science Department, Worcester State University, MA, USA; University of Duhok, Duhok, KRG, Iraq

Abstract: Deep learning techniques like convolutional neural networks (CNN) have not been successfully implemented in brain tumour segmentation due to overfitting issues. The systems currently being used to segment brain tumours are largely manual, which poses limitations regarding the accuracy of results obtained and the amount of time taken to reach the outcome. This research aims to solve the overfitting problem by using drop-out for regularization, so that surgeons can benefit from this technique for timely segmentation of brain tumours. The proposed system is the Enhanced Classification Loss Function (ECLF) algorithm, which uses a deep convolutional neural network for segmenting brain tumours into four classes, namely necrosis, edema, non-enhancing core, and enhancing core. The proposed taxonomy consists of three steps. The first step extracts valuable features from magnetic resonance image (MRI) data. Secondly, random samples are chosen from the im-age patches available for the training dataset. In the third step, the patches are used to train CNNs. Finally, the results are documented to discover the automated segmentation of different tumour classes. The results showed that by using a drop-out layer for regularization, the results provided

a significant accuracy improvement to 94% against the current accuracy of 78%. The proposed system also reduces the processing time by 3-4 frames per second compared to the state-of-the-art method. The proposed ECLF focused on the overfitting problem to improve the accuracy of early detected cancer and reduce the processing time.

A Review of UAV Visual Detection and Tracking Methods

Raed Abu Zitar, Mohammad Al-Betar, Sofian Kassaymeh
Sorbonne Center of Artificial Intelligence, Sorbonne University-Abu Dhabi, Abu Dhabi, United Arab Emirates;
Artificial Intelligence Research Center (AIRC), Ajman University, Ajman, United Arab Emirates;
Artificial Intelligence Research Center (AIRC), College of Engineering and Information Technology,
Ajman University, Ajman, United Arab Emirates

Abstract: This paper presents a review of techniques used for the detection and tracking of UAVs or drones. There are different techniques that depend on collecting measurements of the position, velocity, and image of the UAV and then using them in detection and tracking. Hybrid detection techniques are also presented. The paper is a quick reference for wide spectrum of methods that are used in the drone detection process.

Window-based Lane Localization and Characterization for Self-driving Vehicles

Jiajun Chen, Mark Paulik
Department of ECE & CS, University of Detroit Mercy, Michigan, USA

Abstract: A vision-based lane detection system is an important component of autonomous vehicles. In this work we present a robust road-lane detection and tracking algorithm based on the histogram distribution of lane-tracking windows. In our proposed method, raw images are first normalized, enhanced and segmented using filtering, thresholding, and morphological techniques. Then, these processed images are mapped to a bird's-eye-view perspective for lane tracking and fitting using sliding windows and a polynomial RANSAC algorithm. Preliminary simulation results demonstrate that, when compared to a more traditional Hough-Transform-based approach, the developed algorithm is more reliable and robust on highway and urban roads despite various environmental challenges such as fragmentary road-lane markings, lower resolution images, high degree curves, and strong noise interference.

Face Recognition Using Fuzzy Inference System

Krysel Altarejos, Tracy Pham, Ruting Jia
Electrical and Computer Engineering Department,
California State University Northridge, Northridge, California, USA

Abstract: This paper covers the theory, design, and simulation of facial recognition with fuzzy inference system. The simulation is designed to recognize an individual from a diverse group of people and ultimately find the best matching picture from a dataset. A large group of students with varying facial expressions is collected in order to create a dataset that is split into training and testing pictures so that the simulation can be properly tested throughout the experiment. Facial landmarks are applied in order to retrieve different coordinate points throughout the face so that the facial feature's relative distances and areas could be calculated for each individual and be fed into the fuzzy system. The fuzzy inference system outputs a value that is used to determine which test picture matches best with the training picture that is initially inputted. The dimensions and the brightness of each picture in the dataset is optimized in order to improve the results. After proper modifications, the facial landmarks are working as expected and the newly optimized results improved with an 80% accuracy rate.

Handwritten Text Recognition Using Fuzzy Inference System

Aqibur Rahman, Juan Carlos Atilano, Ruting Jia
Department of Electrical and Computer Engineering,
California State University Northridge, Northridge, California, USA

Abstract: This paper discusses the design and simulation of a lowercase handwritten text recognition algorithm that utilizes a fuzzy inference system. The algorithm allows a user to take photos of digital handwritten letters. Then it automatically detects features of the input image based on Maximally Stable Extremal Regions (MSER) algorithm. Next it filters out non-text features based on morphological differences between text and non-text regions to determine the stroke width of the letter. The image is then processed into different segments for calculations. The segmented image is utilized and implemented in a carefully designed fuzzy inference system to determine what each letter is in any given image.

Resolving User(s) of Interest in the Context of a Multi-Party Collaborative HRI Setting Using Active Speaker and Facing State Estimation

Shuvo Kumar Paul, Pourya Hoseini, Mircea Nicolescu, Monica Nicolescu
Department of Computer Science and Engineering, University of Nevada, Reno, Nevada, USA

Abstract: This paper presents an approach to real-time detection of the User of Interest (UOI) in collaborative Human Robot Interaction (HRI) setting. The presented method tries to detect the UOI by 1) estimating whether the user is facing the robot (or camera), and 2) detecting the active speakers. A dataset was generated by extracting facial landmarks of participants in predefined settings. Several machine learning algorithms were trained to estimate the facing state, and the distance between the landmarks on the lips was compared against a threshold value to determine the active speaker. The paper further presents an evaluation of the proposed methods for active speaker detection and facing state estimation along with demonstrating its real-time application for UOI estimation.

Optimizing Minkowski Distance for Pattern Classification and Regression

Lewis Westfall, Teryn Cha, Sung-Hyuk Cha
Computer Science Department, Pace University, New York, USA;
Computer Sciences, Essex County College, Newark, New Jersey, USA

Abstract: Distance measures are used in many algorithms for machine learning and classification, such as k Nearest Neighbor (k-NN), k-means, k-medoids, etc. Finding an appropriate distance measure is important because of the significance of these algorithms. The most widely used classical distance measures include Manhattan L1 and Euclidean L2 distance measures. Minkowski generalized these measures into the Minkowski metric, Lp for $p \geq 1$. This research searches for the optimal p values on several public data sets, such as the MNIST and Wine Cultivars databases, and intriguingly, the optimal p values found are neither $p = 1$ (Manhattan) nor $p = 2$ (Euclidean). Moreover, p should be greater than or equal to one in order for the distance to be metric, but we found at least one data set where the optimal value p is less than one. These experimental results and reasons for the better performance using the non-metric distance measures are presented.

A Kalman Filter Approach to Eye-tracking Cursor Control for a Wheelchair Interface

Matthew Teague, Joseph Lynch, Nick Bacskai, Ethan Blodgett, Eleanett Perez Cervera, Bahram Shafai, Thomas Consi
Department of Electrical & Computer Engineering, Northeastern University, Boston, Massachusetts, USA

Abstract: Quadriplegics are limited in how they are able to interface with their wheelchairs. We describe an eye-tracking wheelchair interface that allows users to control a cursor with their gaze and activate buttons corresponding to wheelchair commands. The cursor provides the user with instant feedback, making our interface responsive and intuitive to use. Gaze direction data is derived from landmarks produced by a face alignment algorithm and fed into a Kalman filter. Using a Kalman filter approach greatly reduces the noise in the gaze direction data and allows the cursor to move naturally. A webcam is used to capture gaze data, making our interface unintrusive and cost effective.

Automatic Classification of White Blood Cell Images using Convolutional Neural Network

Rabia Asghar, Arslan Shaukat, Usman Akram, Rimsha Tariq
National University of Sciences and Technology (NUST), Islamabad, Pakistan

Abstract: Human immune system contains white blood cells (WBC) that are good indicator of many diseases like bacterial infections, AIDS, cancer, spleen, etc. White blood cells have been sub classified into four types: monocytes, lymphocytes, eosinophils, and neutrophils on the basis of their nucleus, shape and cytoplasm. Traditionally in laboratories, pathologists and hematologists analyze these blood cells through microscope and then classify them manually. This manual process takes more time and increases the chance of human error. Hence, there is a need to automate this process. In this paper, first we have used different CNN pre-train models such as ResNet-50, InceptionV3, VGG16 and MobileNetV2 to automatically classify the white blood cells. These pre-train models are applied on Kaggle dataset of microscopic images. Although we achieved reasonable accuracy ranging between 92 to 95%, still there is need to enhance the performance. Hence, inspired by these architectures, a framework has been proposed to automatically categorize the four kinds of white blood cells with increased accuracy. The aim is to develop a convolution neural network (CNN) based classification system with decent generalization ability. The proposed CNN model has been tested on white blood cells images from Kaggle and LISC datasets. Accuracy achieved is 99.57% and 98.67% for both datasets respectively. Our proposed convolutional neural network-based model provides competitive performance as compared to previous results reported in literature.

ERGOGENIUS: An Integrative Framework for Preventing Work-related Upper Limbs Musculoskeletal Disorders through Image Analysis Techniques

Cedrick Bamba Nsimba, Elie Fiogbe, Reijane Salazar Costa
Computing Department, Federal University of Sao Carlos, Sao Carlos, SP, Brazil;
Physical Therapy Department, Federal University of Sao Carlos, Sao Carlos, SP, Brazil;
Educational Department, Federal University of Sao Carlos, Sao Carlos, SP, Brazil

Abstract: Work-related Upper limbs musculoskeletal disorders (WUMSD) are the most prevalent occupational illness worldwide. They represent a major source of disability and lost work time, therefore resulting in an increased loss in productivity. Their etiology is multifactorial with biomechanical and work organizational aspects being important contributing factors. However, most of the currently available interventions address WUMSDs through improving a given factor singly. This paper aims to present and evaluate the ERGOGENIUS framework, a system which aims to prevent WUMSDs by managing both biomechanical and work space risk factors through image analysis techniques. Twenty employees, computer users in the workplace, from an information technology department were randomly divided in a control (CG) and feedback (FG) groups. Both groups performed an intervention including work space ergonomic evaluation, movement pattern training as well as 3-minute physical training sessions after each working hour through the ERGOGENIUS. This system allows continuous monitoring of user's specific points such as eyes, coracoid processes, ear lobes, elbows and wrists while the user is working or during the exercise sessions. However, for the CG,

the continuous monitoring and posture tracking system were disabled. After a 4-month intervention, when compared to the CG, the FG showed a significant reduction with large effect size of low back (p= .008; d = 0.90) and wrist pain (p= .002; d = 1.02), and number of production days lost (p= .04; d = 0.90) as well as a significant increase with large effect size of handgrip strength (p= .03; d = 0.87). The continuous monitoring allowed by the ERGOGENIUS was effective to prevent the main symptoms of WUMSDs as well as to improve workers' productivity.

Video Dehazing Using Encoder-Decoder Based Deep Neural Networks

Hsing-Che Wu, Jin-Jang Leou
Department of Computer Science and Information Engineering,
National Chung Cheng University, Chiayi, Taiwan, Republic of China

Abstract: In this study, an encoder-decoder based deep neural network for video dehazing is proposed. Performing video dehazing via early fusion, five consecutive (stacked) hazy frames containing spatial and temporal information are fed into the proposed network, and the proposed network will generate the corresponding transmission map of the third (central) frame. The atmospheric light of a frame is estimated among adjacent frames. Skip connections between the corresponding layers at the encoder and the decoder are employed. Batch normalization and two activation functions, namely, leaky rectified linear unit (LeakyReLU) and rectified linear unit (ReLU) are also employed. Finally, the estimated transmission map and atmospheric light of each hazy frame are used together to generate the corresponding dehazed frame based on the atmospheric scattering model. Based on the experimental results obtained in this study, in terms of average PSNR (dB) and average SSIM as well as subjective evaluation, the performance of the proposed approach is better than those of four comparison approaches.

Multi-focus Image Fusion Using Deep Neural Networks

Meng-Che Chung, Jin-Jang Leou
Department of Computer Science and Information Engineering, National
Chung Cheng University, Chiayi, Taiwan, Republic of China

Abstract: In this study, a multi-focus image fusion approach using deep neural networks is proposed, in which focused and unfocused regions are determined by a patch-based deep neural network. The proposed approach contains four main stages. (1) Multi-focus images are transformed into gray-level images and then Laplacian-filtered images. (2) Each Laplacian-filtered image is divided into 16×16 patches in a pixel-wise overlapping manner. The co-location patches of two Laplacian-filtered images are concatenated vertically into 32×16 macro-patches, which are fed into the proposed network to generate initial focus maps. (3) Decision maps are obtained by binary classification and post-processing. (4) Based on the decision maps, each pre-fused image and the final fused image are generated by pixel-based image fusion. Based on the experimental results obtained in this study, in terms of four objective performance metrics, namely, $\boldsymbol{Q^{AB/F}}$, $\boldsymbol{Q_{NMI}}$, $\boldsymbol{Q_{NCIE}}$, $\boldsymbol{Q_Y}$, and subjective evaluation, the overall performance of the proposed approach is better than those of five comparison approaches.

Optimizing Binary Similarity Measures for Pattern Classification and Regression

Lewis Westfall, Teryn Cha, Sung-Hyuk Cha
Computer Science Department, Pace University, New York, USA;
Computer Sciences, Essex County College, Newark, New Jersey, USA

Abstract: Machine learning, classification, regression, and clustering algorithms, such as k Nearest Neighbor (k-NN), k-means, k-medoids, etc. use similarity measures to compare binary records. Because of the significance of these algorithms, it is important to find an appropriate similarity measure. The most widely used classical similarity measures include Hamming, Tanimoto-Jaccard, and Sørensen-Dice. These last two measures can be generalized using a parameter, p, to specify the weight given to the intersection of the records. This research searches for the optimal p values on the public MNIST data set. Interestingly, the optimal

p values found are neither p = 1 (Tanimoto-Jaccard), nor p = 2 (Sørense-Dice). Moreover, we found that the optimal value p is less than one. These experimental results of the distance measures are presented.

Pothole Detection from Dash Camera Images using YOLOv5

Ashit Patel, Lei Huang, Gustavo Vejerano
Department of ECE, Loyola Marymount University, Los Angeles, California, USA

Abstract: In this paper, we propose a new solution to automatically detect potholes on the road surface from dash camera images using a state-of-the-art deep learning based object detection algorithm, namely, You Only Look Once version 5 (YOLOv5). The dash camera image data were preprocessed and augmented as inputs to Convolutional Neural Network (CNN) models, which are trained to out-put the detected potholes with location bounding boxes. Through transferred learning, different sizes of CNN models with different layer architectures are evaluated in terms of mean Average Precision (mAP) and the number of frames per second (fps) detected. Compared with previous work, experimental results show that our proposed solution using YOLOv5 achieved higher detection accu-racy at faster detection speeds, while enabling tradeoffs between accuracy and speed with three different model size options.

Shadow Removal based on Graph Cut

Shijun Tang, Daniel Wu
Department of Computer & Information Science, Cabrini University, Radnor, Pennsylvania, USA

Abstract: In this paper, shadow removal is provided by using objects extraction and image processing. First, objects extraction has been completed using graph cut. An energy minimization representing the optimal segmentation is obtained via max flow (min cut). Second, aimed to the clutter of segmented section, the further image processing (i.e. background contrast) has been applied to remove the redundancy parts. Experiments demonstrate that good results can be obtained by using the proposed method, even under the complex environments, such as moving object with shadow and undesired edges or boundaries.

Skin Tone Benchmark Dataset for Diabetic Foot Ulcers and Machine Learning to Discover the Salient Features

Masrur Sobhan, Kacie Kalie, Abdullah Al Mamun, Anuradha Godavarty, Ananda Mohan Mondal
Computing and Information Science, Florida International University, Miami, Florida, USA;
Biomedical Engineering, Florida International University, Miami, Florida, USA

Abstract: One in three people with diabetes develops diabetic foot ulcers (DFUs) during their lifetime, which is a major risk factor for amputation and mortality. Optical imaging techniques have been developed to determine the extent of oxygenation to these DFUs, which is vital to assessing DFU wounds. Herein, a machine learning approach was implemented to label different skin tones to eventually correct the tissue oxygenation maps. We developed a skin tone benchmark dataset of 9,000 samples and a machine learning framework to represent this dataset in a reduced dimension of 20 features, which could subsequently be incorporated into our smartphone-based optical imaging de-vice developed for DFU assessments. This allows our technology to be applicable across different racial/ethnic groups of varying skin tones.

Plant Leaf Disease Detection with Deep Learning

Heng Wu, Purushothaman Natarajan, Padma Nimmakayala, Umesh Reddy, Rui Wang, Thanh Le
Department of Mathematics & Computer Science, West Virginia State University, Institute, West Virginia, USA;
Department of Biology, West Virginia State University, Institute, West Virginia, USA;
College of Medicine, University of Florida, Jacksonville, Florida, USA;
Business Information Technology, University of Economics Ho Chi Minh, Ho Chi Minh City, Vietnam

Abstract: Plant diseases are a crucial problem to the world's food supplies. The identification of plant diseases at an early stage is crucial for global health and wellbeing. The traditional diagnosis process involves visual assessment of an individual plant by a pathologist through on-site visits. The shortcomings of traditional methods are obvious, the efficiency and accuracy are very low. Deep learning provides some perfect solution. In particular, the contribution of pattern recognition of convolutional neural networks. Using the convolutional neural network, this paper explores a framework for automatic identification of diseases and insects. Two models of it performs well on both data regression and validation loss on pepper, potato, and tomato of PlantVillage public dataset. Engineers or researchers can find a balance point between efficiency and accuracy under this framework and design a convolutional neural network for plant disease detection that meets your requirements

Compressively Sensed Super-Resolution Magnetic Resonance Image Reconstruction Algorithm Using Convolutional Neural Networks

Krzysztof Malczewski
Faculty of Applied Informatics and Mathematics, Warsaw University of Life Sciences, Warsaw, Poland

Abstract: Magnetic resonance scans have recently struggled with inherent limitations such as spatial resolution and long examination times. A novel, rapid compressively-sensed magnetic resonance high resolution image resolution algorithm is presented in this paper. By combining a highly sparse sampling scheme and the super-resolution reconstruction (SRR) method, this technique addresses these two critical issues. Due to the extremely difficult requirements for the accuracy of diagnostic image registration, the presented technique makes use of image priors, deblurring, parallel imaging, and discrete dense displacement sampling for the deformable human body and motion analysis. Clinical trials as well as phantom-based research have been carried out. It has been demonstrated that the proposed algorithm can improve image spatial resolution while reducing motion artifacts and scan times. Convolutional neural networks (CNNs) perform admirably when used to reconstruct images obtained by compressed-sensing magnetic resonance imaging (CS-MRI). The latter goal of the research was to improve subimages' quality by developing a novel iterative reconstruction method that uses image-based CNNs and k-space correction to preserve original k-space data. CNNs represent a priori information about image spaces in the proposed method. The CNNs are first trained to map zero-filling images onto full-sampled images. The zero-filled part of the k-space data is then recovered. Following that, k-space corrections are used to preserve the original k-space data by replacing unfilled regions with original k-space data. The processes described above are used iteratively.

Sediment Plume and Seagrass Detection in Aerial Imagery of Seagrass Beds

Mehrube Mehrubeoglu, Kirk Cammarata, Hua Zhang, Justin Giessel, Lifford McLauchlan
Department of Engineering, Texas A&M University-Corpus Christi, Texas, USA;
Department of Biology, Texas A&M University-Corpus Christi, Texas, USA;
Department of Electrical Engineering & Computer Science, Texas A&M University-Kingsville, Texas, USA

Abstract: Natural events and human activities, such as river flows, currents, dredging, and aquaculture, result in sedimentary suspended plumes that can potentially affect the health of seagrasses which form a critical benthic habitat for marine life. To properly analyze the effect of sediment plumes on seagrasses, the plume must first be detected and characterized. This study utilized aerial imagery from Redfish Bay, Texas, to detect and characterize human-generated sediment plumes to identify potential impacts to seagrass beds. Seagrass area was also detected and analyzed from the same imagery. Four methods were employed to calculate the plume and seagrass cover as they appear from bird's eye view. The initial work utilized thresholding, k-means clustering, Euclidean distance measure to nearest supervised cluster center, and support vector machines for pixelwise

classification to identify sediment plume and seagrass regions. These methods relied on classical image processing techniques to process and analyze RGB color images. Finally, results in terms of area from each method were compared. The methods show general agreement within 5 percent among the detected pixels that represent the plume and seagrass areas.

The 19th International Conference on Modeling, Simulation & Visualization Methods (MSV'22: July 25-28, 2022, USA)
https://american-cse.org/csce2022/conferences-MSV
https://american-cse.org/csce2022/

Validating an Agent-based Ecological Simulation of the Wolves and Moose of Isle Royale National Park

Charles Eugene Knadler, Jr.
Independent Researcher (formerly: IBM & Utah Valley University), Alpine, Utah, USA

Abstract: An agent-based simulation of the future of the moose-wolf ecology of the Isle Royale National Park is validated by constructing confidence intervals, using the maximum and minimum yearly moose and wolf populations from five hundred simulation runs and evaluating the deviations of five-, and ten-year mean and median simulation projections from the Park's past population data. The deviations of the five-year projections from the Park's data are shown to provide evidence of the existence and impact of changes in environmental and external factors on the Park's populations' morbidity and mortality.

Terrain Generation for Specified RMS Roughness Values

Jeremy Mange
US Army, Ground Vehicle Systems Center, DEVCOM GVSC - Analytical Methods, Warren, Michigan, USA

Abstract: Within the US Army, a particular terrain RMS calculation is commonly used in order to characterize the roughness of a terrain. This characterization is important for a number of applications, most significantly for defining certain mobility standards for ground vehicles, and is thus widely used within modeling and simulation (M&S) applications to measure vehicle performance. As part of a larger uncertainty quantification (UQ) project related to vehicle performance metrics, this paper presents two terrain generation algorithms which produce terrains of specified RMS roughness values. The details of these algorithms are provided, along with a demonstration of their application and results, and a brief discussion of how they will be used as part of the larger UQ project.

Investigation of Hand-Tracking Versus Controllers in Virtual Reality

Devin Shapiro, Leonidas Deligiannidis
School of Computing and Data Science, Wentworth Institute of Technology, Boston, Massachusetts, USA

Abstract: Hand-tracking in Virtual Reality has become more accessible to consumers with the introduction of affordable Head-Mounted Displays equipped with cameras to detect their hands in Virtual Environments. This paper investigates the integrated hand-tracking by comparing it to traditional Virtual Reality controllers. Hand-tracking and controllers were compared based on movement, object grabbing, interacting with a UI, and pushing a button. It was found that hand-tracking is more immersive but less accurate than controllers, which was consistent with previous research on this topic. In addition to these findings, it was also found that controllers currently have more potential than hand-tracking when it comes to performing multiple inputs at once or in quick succession.

Challenges and Opportunities Associated with Technology-Driven Biomechanical Simulations

Zartasha Mustansar, Haider Ali, Lee Margetts, Saad Ahmad Khan, Salma Sherbaz, Rehan Zafar Paracha
Department of Computational Engineering, School of Interdisciplinary Engineering and Sciences, National University of Sciences and Technology (NUST), Islamabad, Pakistan;
Department of Mechanical and Aerospace Engineering, University of Manchester, United Kingdom (UK)

Abstract: This paper presents the principal challenges and opportunities associated with computational biomechanics studies. The underlying cognitive control involved in the process of human motion is inherently complex, dynamic, multidimensional, and highly non-linear. The dynamics produced by the internal and external forces and the body's ability to react to them is biomechanics. With the advent of technology and dawn of machine learning tools, it is believed that machine learning approaches can enable us to embrace this complexity, if we could use three arms of ML i.e. predictive modeling, classification, and dimensionality reduction. Biomechanics, since it deals with motion and mobility come with a huge set of data over time and ML is the fastest and reliable way to handle this. Not only this a lot of other computational (Computer Solvers) and technological advances (Wearable sensors), can let us develop computationally inexpensive frameworks for biomechanics focused studies dealing with a huge amount of data. A lot of misunderstanding arises because of extensive data, standardization of the tools to process this, database for the material property definitions, validation and verification of biomechanical models and analytical tools to model various phenomena using computational and modeling techniques. Study of biomechanics through computational simulations can improve the prevention and treatment of diseases, predict the injury to reduce the risk and hence can strengthen pivotal sectors like sports and lifestyle by. This is why this study picks up the challenges and opportunities associated with biomechanics to help improve, analysis, performance and design for better lifestyle.

Optimal Shipping Strategy of Materials Using Traditional and Newly Developed Method Initial Basic Feasible Solution of Russell and Vogel Method

Chandrasekhar Putcha, Rakeshkumar Mahto, Isaiah Chitica, Aycan Kara, Zana Haddad
Department of Civil and Environmental Engineering, California State University, Fullerton, California, USA;
Computer Engineering Program, California State University, Fullerton, California, USA

Abstract: This paper aims to deal with a consolidated study for optimal value determination using conventional and newly developed methods using the initial basic feasible solution of Russell and Vogel method. Shipping of materials through transports vessels is a complicated process. The methods provided in this paper will be used as an optimal method with fast convergence of transportation problems to help in solving the overall shipping plan. This minimizes the total equivalent uniform annual cost, including shipping cost. Typically, a manager of operations research is responsible for determining the shipping plan for each of the three items: continue shipping exclusively by rail, switch to shipping exclusively by water, and ship by either rail or water. With the help of an Initial Basic Feasible (IBF) solution, optimal solution can be reached without running multiple rounds of iterations and using transportation simplex algorithms. The only requirement is that the destination needs to be met within the constraints of source supply. The different types of IBF Solutions used are North-West Corner Rule, South-East Corner Rule, Row Minima, Column Minima, Matrix Minima, and Vogel's. These concepts are used extensively in various disciplines like Industrial Engineering, Mechanical Engineering, and Civil Engineering.

Mississippi Multimodal Freight Analysis Model

Vivek Sunchu, Canhao Wang, Terrence Conley, Mercy Chebet, Shihmiao Chin, Tzusheng Pei, Ali Abu-El Humos
Department of ECE & CS, Jackson State University, Mississippi, USA;
Transportation Analytics & Decision Sciences Group, Oak Ridge National Laboratory, Tennessee, USA

Abstract: The Mississippi Department of Transportation strives to facilitate a safe and efficient movement of freight within the state due to the growing need to understand what is causing congestion and improve safety measures. The scope of this project is to develop a multimodal freight transportation analysis model for Mississippi for assessing the demand for transportation facilities

and services. This research will enable policymakers, transportation planners, and logistic analysts in various federal, state, and local agencies to assess the demand for transportation facilities and services, energy use, safety, risks, and environmental concerns. Implementation of state-level strategies that support efficient freight movement is therefore essential not only for attracting new industries to move freight into and out of the state, but also for addressing the needs of existing state businesses.This work provides a detailed report and analysis model on shipments to and from Mississippi.

Modeling and Simulation Setup for Kinematics and Compliance Analysis

Sara Boyle
Development Command (DEVCOM) Ground Vehicle System Center (GVSC), Warren, Michigan, USA

Abstract: This paper describes a tool to generate a suspension test rig driver input file to be used with the suspension test rig available with Chrono::Vehicle. The suspension test rig driver file defines how the suspension test rig post actuators should move and the steering should be set to evaluate the kinematics and compliance for a vehicle suspension model.

Temporal Accuracy, Fidelity and Time Representation in Complex System Modeling and Simulation

Jeffrey Wallace
Rocket Technology Systems LLC, Virginia Beach, Virginia, USA

Abstract: Improved realism in modeling and simulation (M&S) efforts is increasingly viewed as a necessity, or requirement, across a wide spectrum of activity in academia, industry, and government. Achieving greater realism depends upon temporal accuracy and fidelity, however, and providing different constructs to represent time with respect to interactions is key. This paper describes different methods to represent system evolution and time passage based on the way modeled entities interact in the real world.

High Quality UAV-Based Geo-Localization Matching Study

Zihan Lin
Adelson Educational Campus, Las Vegas, Nevada, USA

Abstract: The task of object localization from UAV (Unmanned Aerial Vehicle) view is a new research area of image processing. The aim of this paper is to accurately match images captured by UAVs with other images from different devices to determine their locations based on data. This technology can be used for UAV precision express (such as handing out masks without in-person contact during a pandemic) and agriculture big data statistics (such as estimating crop yields in a given area through satellite images). The main aim of this research paper is to realize the object localization from UAV view. When given images or videos from UAV view, the identification system can match these with similar satellite images, use GPS data such as latitude and longitude to locate the object, and verify the localization results. We constructed a real-world multi-source dataset called UAV-166 for UAV-based geo-localization. Due to the high cost of annotating a large quantity of actual data, this paper adopts the weakly supervised object localization to match satellite images with UAV view images under the condition that the accurate coordinates of the target building cannot be obtained. Furthermore, we proposed a multi-scale matching method (MSMM) to match satellite images and UAV perspectives at different levels to achieve higher accuracy. Based on experiments on the UAV dataset, our scheme has the RANK @ 1 accuracy of 69.52% with AP accuracy of 74.43%.

The 28th International Conference on Parallel & Distributed Processing Techniques & Applications (PDPTA'22: July 25-28, 2022, USA) https://american-cse.org/csce2022/conferences-PDPTA https://american-cse.org/csce2022/

Implementing Server-side Direct Memory Transfer Communication for Efficient Interprocess Ray-tracing Queries on High Performance Computing (HPC) Systems

Barry C. White, Abdias I. Santiago-Lugo, R. Harrison Hunter, Reena R. Patel
U.S. Army Corps of Engineers, Engineer Research and Development Center, Vicksburg, Mississippi, USA

Abstract: Large-scale distributed memory applications on the HPC allow for efficient processing of data but require time and precious nodes to load the information. It would be beneficial to allow such software to act as a data processor/server for other applications with direct memory transfer. Such a system has been devel-oped utilizing a distributed memory Geometry Engine and Sensor Instances to generate imagery. The goal of this work is to provide a server-side, direct memory communication transfer using ADIOS2 for communication between the Geome-try Engine and multiple software instances. The Geometry Engine provides meth-ods to perform ray-traced queries. The Sensor Instance makes queries to numeri-cally model and generate sensor-specific imagery. This study discusses this im-plementation and the development of an API.

Puzzles for Comparing Real-time Model Checkers

Mitchell L. Neilsen
Department of Computer Science, Kansas State University, Manhattan, Kansas, USA

Abstract: This paper describes some puzzle problems that can be used to evaluate model checkers. Problem models are encoded using three different verification tools: SPIN, TLA+, and UPPAAL. Characteristics of the model checkers and their modeling languages are described. The time and space taken to verify key properties to solve a few puzzle problems is given to analyze and compare performance and limitations of each model checker. Puzzles are easy to grasp, and they expose the limitations of model checkers and the challenge of limiting the size of the search space through careful design. By model checking puzzles, students gain practical experience in developing simple models, before moving on to create models for larger real-world problems. By comparing performance, they also gain insight into methods that can be used to derive efficient models.

End-to-end Trusted Execution Environment (TEE) with Dynamic Holistic View Based Throughput Maximization Approach

Muhammed Akif Agca
Computer Engineering Department TOBB University of Economics and Technology, TOBB ETU, Ankara, Turkiye Luxembourg Institute of Science and Technology / LIST, Esch-Sur-Alzette, Luxembourg

Abstract: Drastic progress and improvements on classical behavior modelling approaches; such as, cellular automata, chaotic systems, hierarchical block diagram modeling methods enabled to avoid of cumborsomism while adapting the dynamism at massive scale at some extend. However, persisting and ensuring the trust for varying contexts with an E2E trust mechanism require ***dynamic holistic views*** to adapt the dynamism at massive scale with extended data locality to the edges in trusted scalable manner. Initial observations for data exchange over a hybrid-cloud node, instead of cell unit scenario in 5G environment with the trust mechanism is promising to meet zero latency requirement of MEC (Multi-access/Mobile Edge Computing) edge units thanks to the improvements provided via ***memory-centric*** system design paradigms. It shows that data can be transmitted over a hybrid-

cloud node rather than cell units can maximize total system throughput of emerging hybrid-clouds, which have 5/6G connectivity and strong quantum back-end units with the E2E trusted execution environment (TEE) and dynamic holistic views. By that means, it is promising to utilize MEMCA hybrid-cloud as ***massive scale cyber-intelligence system*** within the national security legal constraints with the E2E TEE, which have maximized total system throughput via dynamic holistic views of the observed chaotic context. So that, we can say that efficient utilization of MEMCA hybrid-cloud to national security systems as digital dynamics core mechanism can port the massive chaos in socio-dynamics to massive-growth via the dynamic feedback controller structures and embedded check-points to the available physical locations within the (near) real-time cyber intelligence mechanisms with maximized total system throughput values.

Towards Runtime Support for Domain-Specific Languages of Adaptive and Irregular Applications

Polykarpos Thomadakis, Nikos Chrisochoides
Department of Computer Science, Old Dominion University, Norfolk, Virginia, USA

Abstract: We present performance and ease-of-use improvements to a runtime for Domain Specific Languages of irregular applications. Support for message-driven global address space is integrated with lightweight threads which in combination with fine-grained concurrency improves the system's performance and load balancing in shared and distributed memory. We observe up to 100% difference in performance behavior for different lightweight thread creation strategies. Evaluations on a 1960-core distributed memory machine show that the integration of fine-grained concurrency with the runtime achieves performance improvements of 12% on a seismic wave simulation benchmark, as opposed to 50% degradation with OpenMP. Studies on workload decomposition on the same benchmark showed that over-decomposition on both data and task level produces the best results.

Precise FEM and Cutting Simulation of Deformable Bodies on GPU

Harshvardhan Das, Subodh Kumar
Indian Institute of Technology Delhi, New Delhi, India

Abstract: We propose a parallel GPU FEM solver that supports cutting. It has been designed for virtual surgery simulators. The algorithm works on tetrahedral meshes and focusses on interactive but accurate cutting and deformation computation. We subdivide the tetrahedra sliced by a surgical tool. The data structure is designed for coalesced input/output. We also propose a novel incremental method to further improve the FEM-solver's performance. Since most of the matrix/vector entries are zero in a sparse system, we employ a GPU-friendly compressed filter for these. This allows us to load only non-zero entries after a quick filter. We show, experimentally, that our simulation runs in real-time for large meshes also (with size over 1.5M tetrahedra). We also show that our algorithm is optimized for attaining high scalability, and retains mesh quality during the process. This is the first time that precise deformation and cutting simulation for over one and a half million tetrahedra has been reported to complete in less than 18 ms on a single GPU system. Our worst computation times for a time-step are less than 14ms in cutting for the brain mesh (0.45M tetrahedra), and less than 6ms in the FEM solver on the heart mesh of 1.6M tetrahedra. We compare our result with that of the state-of-the-art solver, MFEM, and demonstrate that our results are significantly improved.

Selective Filtering of Large PDF Document Sets: A Performance Oriented Approach

Glenn Davidson, Shikharesh Majumdar
Chancellor's Professor, Director, Real Time and Distributed Systems Research Centre,
Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada

Abstract: A large data volume requires large storage space and increases the time to perform searches within the data by the user. The research presented in this work in progress paper focuses on filtering of large data sets containing documents in the Portable Document Format (PDF) based on keywords/phrases specified by the user. Initial research on a parallel processing-based data filtering technique is presented. Preliminary experimentation with a prototype of the technique deployed on Apache Spark shows a large reduction in stored data volume typically leads to the lowering of search latencies.

Benchmarking Performance of Deep Learning Model for Material Segmentation on Two HPC Systems

Warren R. Williams, S. Ross Glandon, Luke L. Morris, Jing-Ru C. Cheng
Information Technology Laboratory, US Army Engineer Research and Development Center, Vicksburg, Mississippi, USA;
Department of Computer and Information Science & Engineering, University of Florida, Gainesville, Florida, USA

Abstract: Performance Benchmarking of HPC systems is an ongoing effort that seeks to provide information that will allow for increased performance and improve the job schedulers that manage these systems. We develop a benchmarking tool that utilizes machine learning models and gathers performance data on GPU-accelerated nodes while they perform material segmentation analysis. The benchmark uses a ML model that has been converted from Caffe to PyTorch using the MMdnn toolkit and the MINC-2500 dataset. Performance data is gathered on two ERDC DSRC systems, Onyx and Vulcanite. The data reveals that while Vulcanite has faster model times in a large number of benchmarks, it is also more subject to some environmental factors that can cause performances slower than Onyx. In contrast the model times from Onyx are consistent across benchmarks.

Parallel Computing for Large Data Visualizations: A Review

Ananth N. Ramaseri Chandra, David Apostal
School of EE & CS, University of North Dakota, Grand Forks, North Dakota, USA

Abstract: Scientific visualization of large-scale data provides deeper insights and plays a crucial role in analyzing it. Large-scale data visualization is always a challenging process. Processing technologies have increased the computational power of computers to support parallel visualizations with MPI, OpenMP, and CUDA architectures. Tools and APIs based on the parallel visualization frameworks have produced some significant speedup when compared to serial data visualization techniques. This paper presents a review of some of the visualization toolkits and APIs that are based on the visualization pipeline using parallel visualization framework. We also discuss the types of parallelism, features of the tools (like scalability, performance, concurrency), in the field of parallel visualizations and their application areas.

Shadow Process to Improve Running Speed of a Program

Chong Tian, Noha Hazzazi
Department of Electrical Engineering and Computer Science, Howard University, Washington DC, USA

Abstract: Pipeline is absolutely an indispensable technique for modern fast CPU. With the ever-growing requirement for faster CPU and the end of Moore's law, it becomes more important than before. One major obstacle preventing us from achieving its full potential is pipeline hazards. By introducing extra hardware, this paper explores an innovative way called "shadow process" to overcome pipeline hazards especially control hazards . Shadow process duplicate a small but critical portion of original process

(usually several bytes) when a branch instruction is met. By duplicating critical data and making a different prediction than the original process, shadow process bifurcates the original process with the least cost to thoroughly explore the whole predicting space, therefore, it achieve 100% predicting rate which is better than current stage-of-art predictors who intrinsically have a guessing nature. Shadow process only operates when system has idle CPU cores and it can guarantee the integrity of data without breaking current protection mechanism of operating systems.

A Cloud-Edge Continuum Experimental Methodology Applied to a 5G Core Study

Samuel Rac, Rajarshi Sanyal, Mats Brorsson
Interdisciplinary centre for Security, Reliability, & Trust, University of Luxembourg, Luxembourg;
Proximus Luxembourg SA, Luxembourg

Abstract: There is an increasing interest in extending traditional cloud-native technologies, such as Kubernetes, outside the data center to build a continuum towards the edge and between. However, traditional resource orchestration algorithms do not work well in this case, and it is also difficult to test applications for a heterogeneous cloud infrastructure without actually building it. To address these challenges, we propose a new methodology to aid in deploying, testing, and analyzing the effects of microservice placement and scheduling in a heterogeneous Cloud environment. With this methodology, we can investigate any combination of deployment scenarios and monitor metrics in accordance with the placement of microservices in the cloud-edge continuum. Edge devices may be simulated, but as we use Kubernetes, any device which can be attached to a Kubernetes cluster could be used. In order to demonstrate our methodology, we have applied it to the problem of network function placement of an open-source 5G core implementation.

Billion Scale Tensors: Compression Methods and Parallel Computations

Sudhindra Gopal Krishna, Sridhar Radhakrishnan, Aditya Narasimhan, Richard Veras
School of Computer Science, University of Oklahoma, Norman, Oklahoma, USA

Abstract: Tensors, their representations, and computations are gaining popularity with an increasing number of machine learning algorithms for applications. In this paper, we show several techniques for storing compressed sparse tensors. One of our approaches using compressed binary trees (CBT s) shows that the large sparse tensors occupy nearly 1/3 of the space occupied by the storage technique used in tsparse. We show how tensor multiplication can be performed directly on our proposed structure thereby making it space efficient – allowing us to perform tensor multiplication on extremely large tensors on GPU computing platforms. Our results show that the smaller the block size (up to a point), the larger the size of the representation, and the larger the level of parallelism, and vice versa.

A Link Lifetime-based Min Cost Max Flow Routing in a Multi-Layer Edge Cloud Network Environment

Sayed Chhattan Shah
Department of Information Communication Engineering,
Hankuk University of Foreign Studies, Republic of Korea

Abstract: Edge cloud network consists of several mobile and stationary devices interconnected through multiple local area networking technologies such as Ethernet and Wi-Fi. This network is characterized by multiple layers, overlapping network topologies, diverse source-to-destination paths, and a dynamic environment. The multi-layer edge cloud network needs to fulfill the data transfer requirements of resource-intensive and real-time applications such as intelligent video surveillance. Routing plays an integral role in this scenario. In literature, numerous multipath and maxflow-based routing protocols are proposed to address network congestion, latency, throughput, network utilization, network lifetime, and traffic load balancing problems. However, most of these protocols are designed for a homogenous single-layer network environment and do not consider link lifetimes, link energy, and the application's deadline requirements. In this paper, a link lifetime aware minimum cost k flow-based routing

algorithm is proposed to fulfill the requirements of emerging data-intensive and deadline-oriented applications. The algorithm (a) select multiple paths for simultaneous transmission of application data to reduce data latency, (b) select high capacity paths for deadline-oriented applications to meet deadlines and low power consumption paths for energy-sensitive applications to reduce power consumption, and (c) take decisions based on path lifetime and energy to reduce data latencies and power consumption.

Performance of Linear Hashing and Spiral Hashing

Arockia David Roy Kulandai, Thomas Schwarz
Department of Computer Science, Marquette University, Milwaukee, Wisconsin, USA

Abstract: Linear Hashing is an important ingredient for many key-value stores. Spiral Storage was invented to overcome the poor fringe behavior of Linear Hashing, but after an influential study by Larson, seems to have been discarded. Since almost 50 years have passed, we repeat Larson's comparison with in-memory implementation of both to see whether his verdict still stands. Our study shows that Spiral Storage has slightly better look-up performance, but slightly poorer insert performance.

Cooling Power Waste Evaluation Resulting from Malicious Thermal Measurements in Multicore Processors

Michael Kausch, Andrew Ibrahim, Mostafa Abdelrehim
California State University, Bakersfield, California, USA

Abstract: Security attacks on thermal sensors cause the improper reporting of hardware temperatures. This can lead to a decrease in the lifetime of hardware and also an unnecessary increase in power consumption. These security threats due to the presence of malicious thermal sensors can significantly increase energy consumption of cloud computing data centers e.g. Amazon cloud, Microsoft Azure, etc. This poster paper examines the excess power consumption that results from the improper reporting of thermal readings on a Raspberry Pi which can be used to model cloud computing centers in a small scale fashion.

WORKSHOP on Mathematical Modeling and Problem Solving (MPS)

Chairs: Prof. Masahito Ohue, Prof. Masakazu Sekijima*, Prof. Kazuki Joe***
** Tokyo Institute of Technology, Japan*
***chair of Life Computing and Communication Department, Nara Women's University, Japan*

High-Dimensional Bayesian Optimization with Constraints: Application to Powder Weighing

Shoki Miyagawa, Atsuyoshi Yano, Naoko Sawada, Isamu Ogawa
Information Technology R&D Center, MITSUBISHI Electric Corporation, Kamakura, Japan

Abstract: Bayesian optimization works effectively optimizing parameters in black-box problems. However, this method did not work for high-dimensional parameters in limited trials. Parameters can be efficiently explored by nonlinearly embedding them into a low-dimensional space; however, the constraints cannot be considered. We proposed combining parameter decomposition by introducing disentangled representation learning into nonlinear embedding to consider both known equality and unknown inequality constraints in high-dimensional Bayesian optimization. We applied the proposed method to a powder weighing task as a usage scenario. Based on the experimental results, the proposed method considers the constraints and contributes to reducing the number of trials by approximately 66% compared to manual parameter tuning.

A Method of High-Speed Segmentation for Body Motion According to Likelihood Calculation Parallelizing Based on GP-HSMM

Miho Kawamura, Yuichi Sasaki, Yuichi Nakamura
Information Technology R&D Center, Mitsubishi Electric Corporation, Kamakura Kanagawa, Japan

Abstract: In this study, we propose a faster calculation method for body motion segmentation based on the Gaussian process hidden semi-Markov model (GP-HSMM). To remove bottlenecks in calculating forward probabilities, we parallelize them by shifting rows of the log-likelihood matrix, which is calculated from all combina-tions between the signal patterns learned by the Gaussian process and observed data, and summing up rows and columns incrementally. In the experiment with 2593×11 electromyography (EMGs) and inertial measurement unit (IMUs) time series data, our algorithm was 331 times faster than conventional methods.

On a Fast Computation of the Condition L-curve in the LSMR Method for Solving Ill-conditioned Problems using the DQDS Method

Itsuki Kuboi, Rika Tanaka, Shin-ya Ozawa, Yohsuke Hosoda, Masami Takata, Kinji Kimura, Yoshimasa Nakamura
Nara Women's University, Nara, Japan

Abstract: Let us consider an application of the LSMR method, which is one of the iterative solution methods, to an ill--conditioned linear equation with a large sparse coefficient matrix. The approximate solution of the ill--conditioned equation may change significantly even if the noise contained on the left-hand matrix and the right-hand vector is small. When iterative solution methods are used, the iterative computation should be terminated in the middle to reduce the effect of the noise. Therefore, it is necessary to stop iteration at an appropriate number of times when considering the LSMR method for such a problem. In this paper, we propose a fast computation of the condition ~L--curve, which gives an iteration stopping rule, that combines the relative residual norm and the condition number of a bidiagonal matrix. We adopt the DQDS method for computing singular values to compute the condition number and verify that the DQDS method is faster than the bisection method.

Proposal on Scheduling Method for Sightseeing with Unlimited Ride Tickets

Shiori Higano, Reina Isono, Masami Takata
Nara Women's University, Nara, Japan

Abstract: This study proposes a scheduling method for sightseeing with unlimited ride tickets. An unlimited ride ticket allows a user to ride on and off trains and buses in a specified area multiple times. In the case of sightseeing using an unlimited--ride ticket, the time required for walking, as well as train, and bus rides, must be considered. For this purpose, we proposed an evaluation function. This evaluation function is adaptable to several optimization methods. The experiments confirmed the effectiveness of each optimization method.

Calculation of Spectral Similarity Independent of Measurement Equipment

Ryo Murakami, Hiroshi Shinotsuka, Kenji Nagata, Hideki Yoshikawa, Hayaru Shouno
University of Electro-Communications, Graduate School of Informatics and Engineerings, Chofu, Tokyo, Japan

Abstract: We often compared measured photoelectron spectra with other spectra for material development and quality control in the industry. In particular, X-ray photoelectron spectroscopy is used to detect surface contamination and chemical state changes. However, spectral data has perturbation of measurement devices, e.g., the difference in peak width due to the resolution of the device, and the difference in peak position due to the charging phenomenon in the spectral data. It is difficult to simply measure the distance between the measured spectra. Therefore, it is necessary to develop a method for calculating the similarity between spectra that is independent of the device. To establish a comparing procedure, we introduced a clustering method for spectral data to decouple the measurement perturbation. We designed the clustering method for detecting contamination components and sample heterogeneity. This study proposed an analytical model that separates the photoelectron peaks from the perturbation caused by the measurement device. We applied the method to calculate the similarity between the spectra. As a result, we show the proposed method could detect spectral data included with other components in the analysis of real X-ray photo-electron spectroscopy spectral data of TiO2.

Improved Fitness Functions for 3D Packing Problem

Shigeka Nakajima, Masami Takata
Nara Women's University, Nara, Japan

Abstract: In this study, we improved the fitness functions for packing problems. To solve packing problems, a combination of the bottom, depth, and left (BDL) and local search methods was proposed. In the BDL method, all objects were assigned random numbers. The objects were arranged at the bottom, depth, and left in numerical order. Subsequently, the order of the placed objects was changed using the local search method. The fitness of the combination method was then improved. In the combination method, fitness functions represented the frequency of use, viewability, and capacity. These were combined to evaluate the total fitness function. However, conventional methods do not normalize these fitness functions. Therefore, the effect of a fitness function becomes larger than that of the others. In addition, the correlation between these fitness functions remains unclear. In conventional methods, the sizes of all objects are the same. However, this setting is not realistic. Furthermore, objects that are not stored are not considered. Therefore, in this study, we improved the fitness functions through redefinition and normalization.

A New Ensemble Framework based on MOEA/D

Jiayi Han, Shinya Watanabe
Muroran Institute of Technology, Muroran Hokkaido, Japan

Abstract: Based on the "*No Free Lunch*" theory, only a fixed algorithm for generating new solutions in the original MOEA/D cannot efficiently solve all MOPs. In this paper, we propose a new framework based on MOEA/D named MOEA/D-EF (Ensemble Framework), which can contain a variety of new-solutions generating algorithms (candidate algorithms) with different search capabilities to improve the overall universality of the algorithm. In numerical experiments, we take our original DE variant based on the ideal point and historical information as one of the candidate algorithms for generating new solutions. The numerical experiments show that the new framework has broader universality.

Random Population-based Decomposition Method by Linkage Identification with Non-linearity Minimization on Graph

Rui Zhong, Masaharu Munetomo
Faculty of Information Science and Technology, Hokkaido University, Sapporo, Japan;
Information Initiative Center, Hokkaido University, Sapporo, Japan

Abstract: Cooperative co-evolution (CC) has been employed to solve large-scale optimization problems. In theory, perfect decomposition can exponentially decrease the search space without losing optimization accuracy. However, it is impossible to design a high-accuracy decomposition method with little prior knowledge in practice. High-accuracy decomposition methods often consume massive fitness evaluation times to detect the interactions between variables. The representative algorithms are Linkage Identification and Differential Grouping (LINC-R/DG). Large computational cost and local non-linearity check make these algorithms limited. This paper proposes a novel method that regards decomposition as an optimization problem and designs an objective function based on LINC-R/DG. We mathematically explain the feasibility of our objective function. Experiments show that our decomposition method has broad prospects for solving large-scale problems.

Consideration on Automatic Generation of Injection Attacks using Neural Networks

Takeshi Matsuda, Michio Sonoda
Graduate School of Management Information, Hannan University, Amami, Japan;
Higashi, Matsubara, Osaka, Japan

Abstract: Various countermeasures for injection attacks on web applications have been considered and investigated until now. Nevertheless,new vulnerabilities have been reported and new methods to avoid attack detection continue to be developed. Many tools for finding such vulnerabilities of web applications have been developed, but there are cases where these tools are also exploited by attackers.Considering the current situation of such injection attacks, it is important to develop new attack methods earlier than attackers and prepare for unknown attacks. This study used a neural network to learn the structure of data that can actually be used for OS command injection attacks, and verified whether effective attack data can be generated.

Application of Deep Metric Learning to Early-modern Japanese Printed Character Recognition

Norie Koiso, Yuki Takemoto, Sayaka Iida, Yu Ishikawa, Masami Takata, Kazuki Joe
Nara Women's University, Nara, Japan;
Shiga University, Japan

Abstract: Early-modern books published in the Meiji, Taisho and the first twenty years of Showa era are available in image format in the Digital Collections of the National Diet Library. Although these books are required to be converted into text for information utilization, it is difficult to apply existing OCR to early-modern books. Therefore, research on character recognition methods specifically for early-modern Japanese printed character recognition is underway. Currently, the recognition method using a CNN, which has the best recognition rate, has some character types that are difficult to identify because the regions in the feature space overlap for each character type. In this paper, we propose a character recognition method specialized for early-modern Japanese printed characters using deep metric learning, which can eliminate the overlap of regions.

Translating Early-modern Written Style into Current Colloquial Style in Hoji Shinbun

Honoka Nishikawa, Yuki Takemoto, Sayaka Iida, Yu Ishikawa, Masami Takata, Kaoru Ueda, Kazuki Joe
Nara Women's University, Nara, Japan

Abstract: Even in the Early-modern era, the style of the Japanese language has undergone significant changes due to the movement for unification of speech and writing during the Meiji era, and only a limited number of people can easily read the Japanese language nowadays. It is also known that the translation evaluation index BLEU, which is widely used in interlanguage translation, tends to calculate a higher value than the actual translation accuracy when applied to Japanese. In this paper, we perform experiments for automatic translation of early-modern written style and current colloquial style using neural machine translation. In addition, the validity of the BLEU is discussed, and a valid translation evaluation index is examined for early-modern written style and current colloquial style.

The 21st International Conference on Security & Management (SAM'22: July 25-28, 2022, USA) http://sam.udmercy.edu/sam22 https://american-cse.org/csce2022/

Towards Security Enhancement of Blockchain-based Supply Chain Management

Lamees Al-Alawi, Raiya Al-Busaidi, Abdul Khalique Shaikh
Department of Information Systems, Sultan Qaboos University, Muscat, Oman

Abstract: The cybersecurity of modern systems has dramatically increased attention from both industrial and academia perspectives. In the recent era, the popularity of the blockchain-based system has traditionally been emergent among various industrial's sectors especially in supply chain management due to its streamlined nature. This reveals the importance of the quality aspects from a supply chain management perspective. Many industries realized the importance of having quality systems for supply chain management and logistics. The emergence of blockchain technology has created several potential innovations in handling and tracking business activities over the supply chain processes as specific. This paper shed the light on the blockchain and specifically on a smart contract technology which been used to handle the process of creation, verification and checking data over the supply chain management process. Then, touch upon the area of blockchain cybersecurity in the supply chain context. More and more, since the smart contract handles the transfer of data over different locations, then the security protection should be strong enough to secure the data and the assets from any attacks. Finally, the paper examines the main security attacks that affect the data on the blockchain and propose a solution.

A Review of Encryption Algorithms used in Cloud Computing

Derick M. Rakgoale, Topside E. Mathonsi, Vusumuzi Malele
Tshwane University of Technology, Pretoria, South Africa

Abstract: Cloud computing offers distributed online and on-demand computational services from anywhere in the world. Cloud computing services have grown immensely over the past years especially in the past year due to the Coronavirus pandemic. Cloud computing has changed the working environment and introduced work from work phenomenon which enabled the adoption of technologies to fulfil the new workings including cloud services offerings. The increased Cloud Computing adoption has come with new challenges of data privacy and its integrity in the cloud environment. Previously advanced encryption algorithms failed to reduce the memory space required for Cloud Computing performance thus increased the computational cost. This paper reviews the existing encryption algorithms used in Cloud Computing. In the future, Artificial Neural Networks (ANN) algorithm design will be presented as security solution to ensure data integrity, confidentiality, privacy, and availability of user data in Cloud Computing. Moreover, MATLAB will be used to evaluate the proposed solution, and simulation results will be presented.

A Review of Intrusion Detection Algorithms used in Cloud Computing

Kgothatso A. Kwinana, Topside E. Mathonsi, Vusumuzi Malele
Tshwane University of Technology, Pretoria, South Africa

Abstract: Intrusion detection plays a very important part in information security and is a key technology in detecting various types of attacks. As a result, intrusion detection system (IDS) is becoming a critical piece of the security puzzle in Cloud Computing. Due to the increasing number of attacks and their complexity, protecting the Cloud Computing environments from malicious attacks is incredibly difficult. This review paper presents existing solutions developed for intruder detection in Cloud Computing. Furthermore, we outline the strengths and challenges of these exiting techniques. This comprehensive study serves as a starting point and a guide for everyone interested in exploring research in intruder detection in Cloud Computing.

Evaluation of an Information Security Management System at a Mexican Higher Education Institution

Rusbel Dominguez Dominguez, Omar Flores Laguna, Jair A. del Valle Lopez
Facultad de Ciencias Empresariales y Juridicas, Universidad de Montemorelos, Montemorelos, Mexico

Abstract: The purpose of this research was to know the degree of administrative knowledge, the degree of training of human resources, the degree of commitment of administrators and the degree of effectiveness of the administration for information security risk based on ISO/IEC 27001. The population consisted of 81 subjects (66 administrators and 15 ITD personnel). Those evaluated were employers of the administrative office of the university and also staff of the Information Technology Department (ITD). To make the comparisons, three groups of managers were formed according to classifications of administrative staff, the classification was as follows: (a) first-line manager, (b) middle management and (c) top management. About the results, it can be corroborated that administrative staff with a lower rank have more problems in making the best decisions in relation to the implementation of an ISMS, it should be noted that the first-line manager is the one who has more contact with the students and is the one who is less involved in the implementation of an ISMS. It can also be inferred that the institution's planners are not fully trained in the institution's information security efforts. This in turn prevents the generation of proposals for initiatives to implement an ISMS. With this shortcoming, it is possible that security breaches could be generated.

Cybersecurity requirements for Critical Infrastructures

Choong-Hee Han, Chang-Hee Han
Korea Power Exchange, Korea; Korea Military Academy, Korea

Abstract: A cyber-intrusion incident on critical infrastructure has a very serious national and social impact, which will result in paralysis and confusion of national functions. In order to strengthen cyber safety and prevent cyber-intrusion accidents on critical infrastructure, systematic information protection risks in the direction of strengthening the level of blockade by inflow path of cyber threats and strengthening the capability level of security equipment to appropriately respond to incoming cyber threats management is absolutely necessary. To strengthen the cyber safety of critical infrastructure, first, it is necessary to establish a new BDLA (Blockade and Defense Level Analysis)-based information protection risk assessment standard. Second, it should be mandatory to restrict access of overseas IPs to web-based information systems operated by critical infrastructure. Third, it should be mandatory to strengthen the level of information protection of private participants who access the information system of critical infrastructure. Fourth, cyber threat response efficiency should be improved by standardizing the operation sequence of security equipment in critical infrastructure.

Effects of the COVID-19 Pandemic on Organizations' Security Policies

Kathryn Southard, Lila Rajabion
State University of New York (SUNY), Empire State College, New York, USA

Abstract: Before the COVID-19 pandemic, most organizations did not have work-from-home (WFH) models of employment or contingency plans for abrupt transitions to remote work. After the initial pandemic response, businesses reported viewing WFH models more positively, increasing the chances of moving towards more flexible work models post-pandemic. Executives also highlighted the gaps and vulnerabilities in their current security policies, allowing cybercriminals to take advantage of the volatile business environment. In order to maintain business continuity and protect their assets in the future, businesses must update security planning and procedures to include unexpected upheaval in their workforce and provide more flexible work models for their employees.

A Bayesian Deep Learning Regularization for Detecting Zero-Day Attacks

Raaid Wilson, Farid Shirazi
Data Science and Analytics, Toronto Metropolitan University, Toronto, Ontario, Canada;
Ted Roger School of Information Technology Management,
Toronto Metropolitan University, Toronto, Ontario, Canada

Abstract: Intrusion detection systems are part of cyber-security measures determining if a com-puter system has been breached. Many machine learning approaches have been explored to detect different security attacks. However, most of these experiments have not tested the model under zero-day attack circumstances. For example, a zero-day attack would comprise attack patterns for which the models under question have not yet been trained. This study aims to evaluate the effects of Bayesian regularization on a deep neural network for detecting zero-day attacks. By deploying Deep Neural Network (DNN) and Bayesian Neural Network (BNN) on the dataset, our experiment shows that the regular-test error rates of the Bayesian model are higher than those of the other deep neural network models.

Autonomous Penetration Testing System Command Specification

Jack Hance, Jeremy Straub
Department of Computer Science, North Dakota State University, North Dakota, USA

Abstract: Future cyberattacks will likely be distributed and loosely coupled. While attackers will undoubtably devise numerous attack mechanisms, distributed cyberattacks can be modeled by using the Blackboard Architecture. The Blackboard Architecture allows a single or distributed collection of artificial intelligence-powered command node to control a large army of seek nodes, at-tack nodes, and verification nodes. Each type of node executes a specific task. Seek nodes look for end points to target. Attack nodes attack those targets, and verification nodes audit the success of the attack. Attack nodes are designed to execute arbitrary scripts with different requirements based on instructions from their caller. This paper defines a preliminary command specification. This specification is intentionally minimal, with only the absolutely necessary in-formation included. It is defined to allow some commands to fit in a single packet while also being human readable.

Secure Ad Log Storage using Integrated Blockchain and IPFS

Nahid Ebrahimi Majd, Yash Shah
Department of Computer Science and Information Systems,
California State University San Marcos, California, USA

Abstract: Digital advertising, particularly ads on video streams, has had a rapid growth over the last few years. In this research, we propose an integrated blockchain and IPFS framework to securely store, share and tamper-proof ad logs for DVEO, which is a leader in delivering videos and ads over IP. To maintain the scalability and efficiency, we do not directly store the logs on the blockchain. Instead, we utilize IPFS to store logs. Then we hash the log using the IPFS Merkle DAG structure and store the hash on the blockchain. We designed a smart contract to store and access these hashes on the blockchain. The computational complexity of this smart contract to process, store, and access logs is low. Our framework is a scalable, cost-effective, and tamper-proof system adjusted for DVEO and grants access only to the authorized parties. The proposed idea can be adopted to securely store and share any type of sensitive data and guarantee that the data remains immutable.

Consideration of the Use of Quantum Fuzzing for Defense Applications

Dominic Rosch-Grace, Jeremy Straub
Department of Computer Science, North Dakota State University, North Dakota, USA

Abstract: Quantum noise can be a key obstacle for many uses of quantum computers. It results from the environment that the quantum computers operate in and the environmental factors that surround the qubit containment spaces. This noise causes qubits to suffer from decoherence and loose information, which can potentially lead to erroneous outputs in quantum computations. This serves as one of the largest obstacles to the large-scale use of quantum computers. Despite the problems associated with quantum noise, it can also be harnessed as a tool. One such use is for fuzzing, a computing technique that bombards software with numerous randomized input values to test it for security flaws and bugs. Fuzzing automates the process of searching for potential software security flaws and may catch issues missed by human testers. This paper discusses the potential use of the inherently random nature of quantum noise to generate fuzz values and their use for defense applications. It discusses how this approach may offer performance beyond modern classical computing-based fuzzing techniques. It also discusses the use of a fuzzing-like approach for enhancing artificial intelligence techniques performance. Notably, these proposed uses of quantum computing noise for fuzzing make use of an otherwise negative aspect of quantum computing and potentially reduce the cost of fuzzing input value generation while enhancing system performance.

Quantifying the Milestones of Cyber Vulnerabilities

Michael Lidestri, Neil C. Rowe
U.S. Naval Postgraduate School, USA

Abstract: We have developed quantitative models of the durations between the key events in the lifetime of a cyber vulnerability: disclosure time from creation time, path time from disclosure time, exploitation time from disclosure time, and patch time from creation time. Our analysis was based on data from the National Vulnerabilities Database of NIST with additional information from software vendors and the Exploit Database. We observed a significant slowing of patching five years after the release date. We found the cumulative distribution fit well to a Weibull distribution. The duration between disclosure and patching was a symmetric sigmoid curve about zero. We also examined the effects of severity of a vulnerability and operating system under which it occurs.

Advance-Fee Scam Email Classification using Machine Learning

Nahid Ebrahimi Majd, Jorge Sanchez
Department of Computer Science and Information Systems,
California State University San Marcos, California, USA

Abstract: This study investigates the effectives of supervised machine learning algorithms at classifying Advance-Fee fraud email messages, also known as 419 scams. Five models were developed using five popular supervised machine learning algorithms: Naive Bayes, Support Vector Machine, Multilayer Perceptron, Logistic Regression, and Random Forest. The models investigated in this study were compared to models that target general spam. Results showed improvements in classification accuracy of Advance-Fee scams over general spam for all the investigated models. In the case of Logistic Regression, targeting Advance-Fee scam messages showed an accuracy score of 99.1 %, a more than 4 % improvement over Logistic Regression models targeting general spam messages. These findings show that Advance-Fee scam emails should be targeted using machine learning models that are specifically trained for such messages. These findings also imply that targeting specific types of spam is more effective than targeting many types of spam with a single model.

Blockchain Project Workflow Execution for Trustless Operations

Samuel Ashaolu, Lei Chen
Georgia Southern University Statesboro, Georgia, USA

Abstract: In our world today, organizations are adopting various means and pathways to collaborate strategically with other organizations towards boosting their productivity, expanding their business growth, and attaining new frontiers in innovation. One critical challenge has inhibited this process in service collaboration amongst various chains of businesses which is the lack of trust amongst organizations. A forthcoming approach to mitigate these issues is by leveraging the applications of blockchain technology made possible by embedding business execution workflow with smart contracts. This approach can drastically improve transparency, accountability as well as trust decentralization amongst corporations. Blockchain technology is able to provide an immutable and auditable distributed architecture capable of supporting a distributed workflow management system which is the scope of this paper. Overall, it enables the fast automation of business processes and workflows using smart contracts to enforce project agreements across the network. This would bolster collaboration between trustless entities without a dependency on third-party involvement. In general, the existing blockchain-based workflow management systems focus on workflow coordination. This paper thus proposes to illustrate an information workflow mechanism for inter-organizational collaboration enforced via smart contract and deployed on the blockchain network as well as a mini implementation of this process on a Solidity IDE programming platform and associated results. Keywords: Workflow management; Blockchain; Smart Contract; Ethereum; Trustless operations; project automation

NTRU Encryption Algorithm to Enhance Secure Data Transmission for SCADA Systems

Sheng-Tzong Cheng, Po Jung Lai, Hsin Hung Lin, Jia-Ling Li, Gwo-Jiun Horng
Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan;
Regional Industry Service Division, Institute of Information Industry, Nantou, Taiwan;
Department of Computer Science and Information Engineering,
Southern Taiwan University of Science and Technology, Tainan, Taiwan

Abstract: The supervisory control and data acquisition (SCADA) system was designed as a monolithic circuit in the past where a circuit had no need to connect to other circuits. The demand for security enhancement on the SCADA systems is increasing dramatically. In this paper, we study the new enhancement of the SCADA system in which the public key cryptosystem is taken into account for the external communication. We survey several encryption algorithms to find a suitable solution that requires less computation time. We exploit the NTRU encryption algorithm to enhance data security. We consider the real-time SCADA system with the Modbus protocol. We also analyze the performance of the SCADA Configuration Change Management mode.

Security Applications Track Controlling User Fields of View for Safe Smartphone Walking

Kenichi Takahashi, Takao Kawamura
Tottori University, Tottori, Japan

Abstract: Based on the rapid spread of smartphones, accidents that occur when users operate smartphones while walking have become increasingly common. However, it is not impractical to prohibit the use of smartphones while walking owing to their convenience. In this study, we aimed to reduce the risk of using a smartphone while walking without prohibiting smartphone walking. The posture of a smartphone user is typically leaned forward compared to without a smartphone. This makes it difficult for users to see surrounding obstacles because their field of view is narrowed, which can lead to accidents. Therefore, we propose a method that broadens user fields of view by controlling the smartphone holding angle. Experimental results demonstrate that our proposed interface successfully broadens user fields of view.

Space Efficient Secret Sharing Using Repeatable Random Sequence Generators

Arvind Srinivasan, Chien-Chung Chan
SplitByte Inc., Los Gatos, California, USA;
University of Akron, Akron, Ohio, USA

Abstract: We present a new secret sharing algorithm that provides the storage efficiency of an Information Dispersal Algorithm (IDA) while providing perfect secret sharing. We achieve this by mixing the input message with random bytes generated using the Repeatable Random Sequence Generator (RRSG). We also use the bytes from the RRSG to provide random polynomial evaluation points in addition to using the bytes for choosing all the polynomial coefficients and optionally compute the polynomials on random isomorphic fields rather than a single fixed field.

A New Approach for Deep Analysis and Prediction of Smart Grid Stability and Theft Using Machine Learning

Hisham Albataineh, Viswas Kanumuri, Mais Nijim, Avdesh Mishra
Department of Physics and Geosciences, Texas A&M University-Kingsville, Texas, USA;
Department of Electrical Engineering & Computer Science, Texas A&M University-Kingsville, Texas, USA

Abstract: N/A

Improving Security in Smart Cities using Adaptive Age Invariant Face Recognition Model

Kennedy Okokpujie, Okokpujie Imhade P., Agoha Benedict, Osemwegie Omoruyi, Odumuyiwa Ayomikun
Covenant University, Ota, Ogun State, Nigeria;
Department of Mechanical and Mechatronics Engineering, Afe Babalola University, Ado Ekiti, Nigeria;
Department of Mechanical and Industrial Engineering Technology,
University of Johannesburg, Johannesburg, South Africa

Abstract: The deployment of smart technology in cities is usually welcomed as the solution to numerous modern urban challenges such as environmental protection, waste management, and transportation. In most instants, the security subjects (crime detention and prevention) are ignored. Furthermore, researchers rarely deliberate or design and develop prototypes on how new smart city surveillance security cameras might disrupt the customary face detection and recognition system, which automates updating her database without human intervention, which this paper aims to. Hence, it improves security in smart cities using the adaptive age invariant face recognition (AIFR) model. A non-intrusive AIFR model is developed using a convolutional neural network and transfer learning techniques; embedded in the surveillance cameras. The cameras captured the faces of the city dwellers at regular intervals and updated its database. On testing the model, it was found to be capable of enhancing security by detecting, identifying, and updating the database of the smart city residents and visitors. It could also alert the appropriate authorities of criminals or missing persons even with considerable age intra-class variation.

Identity Management for Intelligent Energy Devices Supporting Balance-Responsible Virtual Energy Communities

Marius Stuebs, Kristi Kola, Wolfgang Renz
University of Hamburg, Germany;
HAW Hamburg, Germany

Abstract: The electricity supply and distribution of the future will be controlled digitally. Market platforms for Virtual Power Plants (VPPs) are emerging and compete for dominance in the electricity market, thus not caring for inter-operability. To prevent a monopoly, an approach is to enforce a decentralized, inter-operable infrastructure that enables the formation of self-organizing energy communities beyond local bounds. Such an infrastructure must be capable of supporting a multitude of intelligent energy devices (IEDs) such as electric vehicles (EVs). A secure operation of this critical infrastructure requires to maintain cyber security that guarantees accountability and authentication between the IEDs, to ensure the reliability of the measurements and control commands. In addition, the infrastructure needs to enable a fair electricity marketplace to enforce the continuous balance in the physical grid and take into account the physical safety limits of power lines and substations. In this paper we show how identity management (IDM) and Virtual Energy Communities (VEC) will contribute to solve these challenges. The requirements for an IDM architecture are modeled to enable trust establishment between the IEDs and smart grid services. Using the smart grid architecture model (SGAM) we study EV charging within balance-responsible VEC. The SGAM modeling shows the feasibility of the proposed IDM-based solution to support VEC-based balance.

Fusion of Convolutional Neural Networks for Malware Detection

Yassine Belkhouche
Department of Computer Science, Missouri State University, Springfield, Missouri, USA

Abstract: This paper establishes a convolutional neural network (CNN) decision-level fusion scheme for malware detection. The proposed fusion technique combines the decision made by multiple CNNs into a final decision. We designed and trained several convolutional neural networks using features extracted from malware files. We trained each CNN with different parameters. Two fusion rules are used to combine the individual decisions: majority voting and weighted voting. We evaluated the performance of our method using the EMBER dataset, a publicly available dataset for malware classification. We demonstrated that the fused results outperform the individual decisions in all cases.

LS-AODV: A Routing Protocol Based on Lightweight Cryptographic Techniques for a Network of Nano Drones

Cody W. Vernon, Preetha Thulasiraman
Department of Electrical& Computer Engineering, Naval Postgraduate School, Monterey, California, USA

Abstract: The deployment of a Flying Ad Hoc Network (FANET) of nano drones on the battlefield comes with specific performance and security issues. This paper provides a novel approach to address the performance and security concerns faced by FANET routing protocols, specifically the Ad Hoc On Demand Distance Vector (AODV) protocol. We propose and develop the Lightweight Secure Ad Hoc On-Demand Distance Vector (LS-AODV) algorithm which uses a lightweight stream cipher, Trivium, to encrypt routing control packets, providing confidentiality. The scheme also uses Chaskey-12 based message authentication codes (MACs) to guarantee the authenticity and integrity of control packets. We use a network simulator, NS-3, to compare LS-AODV against two benchmark routing protocols, AODV and the Optimized Link State Routing (OLSR) protocol. We gauge network performance and security benefits of LS-AODV under non-threat and threat (black hole attacks) scenarios. Our results show that LS-AODV provides secure communications with minimal network performance degradation and outperforms the standard AODV protocol when the network is under attack.

Evaluation of User Perception on Biometric Fingerprint System

Jones Yeboah, Victor Adewopo, Sylvia Azumah, Izunna Okpala
University of Cincinnati, USA

Abstract: Biometric systems involve security assurance to make our system highly secured and robust. Nowadays, biometric technology has been fixed into new systems with the aim of enforcing strong privacy and security. Several innovative system have been introduced, and most of them have biometrics installed to protect military bases, banking machines, and other sophisticated systems, such as online tracking systems. Businesses can now focus on their core functions and feel confident about their data security. Despite the benefits and enhancements in security that biometrics offer, there are also some vulnerabilities. This study aimed to investigate the biometric vulnerabilities in a healthcare facility and propose possible countermeasures for biometric system vulnerabilities.

A Survey of the Insider Threat Mitigated Systems Using Blockchain

Wonseok Yoon, Hangbae Chang
Department of Security Convergence, Chung-Ang University, Seoul, South Korea;
Department of Industrial Security, Chung-Ang University, Seoul, South Korea

Abstract: Insider threats are now a major problem within the organization of industry. It is an insider inside a company that leaks sensitive information of the company, such as trade secrets, which can cause huge losses to the company. However, if an internal malicious user gains legitimate rights and deletes, destroys, or modifies the activity history, the system does not consider the behavior to be malicious, so it is difficult to detect or track the malicious insider. In this paper survey on using blockchain, a distributed storage technology in which data is transparently managed and difficult to tempering, to mitigate threats from the inside of organizations. Afterwards, based on previous research results, we analyze common limitations and provide guidelines for future research directions.

Using Data Lineage for Evaluating Damage to Critical Infrastructure Data

Ian Moncur, Brajendra Panda
Computer Science & Computer Engineering Department, University of Arkansas, Fayetteville, Arkansas, USA

Abstract: Rapid and accurate damage assessment is crucial to minimize downtime in critical infrastructure. Fast and consistent techniques are vital to preventing damage from spreading while also minimizing the impact of the damage. To stop the spread of damage, once detected, the entire system is shut down until accurate damage assessment is performed. In critical infrastructure systems, this is unacceptable. The goal of this research is to present a novel model that uses data lineage with the goal of providing fast and accurate damage assessment. In function, this model operates as a directed graph, with the vertices being data items and edges representing dependencies.

Security Algorithms for Identifying Malicious E-mail

Woo Young Park, Sang Hyun Kim, Duy-Son Vu, Chang Han Song, Hee Soo Jung, Hyeon Jo
RealSecu, South Korea

Abstract: The majority of cyberattacks are carried out via e-mail. Existing security solutions blocked malicious mail by forming a blacklist based on the signature. But they are vulnerable to zero-day attacks and have a cold-start problem. This work proposes a new security algorithm using e-mail communication information and a simple mail transfer code (SMTP). The algorithm aims to identify and process malicious mail. It investigates the HELO domain, session domain, pointer (PTR) record, sender policy framework (SPF) record, sending domain, receiving domain, and SMTP reply code. Through this, this algorithm classifies malicious mail into private mail, spear-phishing, and impersonation mail. The performance of the proposed algorithm was verified by introducing it to the three organizations. The result showed that the algorithm could identi-fy the new types of malicious e-mails that had not been detected by existing solutions.

Study on Content-based Information Security Classification using Information Network

Giwan Hong, Hangbae Chang
Department of Security Convergence, Chung-Ang University, Seoul, Korea;
Department of Industrial Security, Chung-Ang University, Seoul, Korea

Abstract: As technological hegemony intensifies, corporates' technological competitiveness is becoming more important than ever. To protect technological competitiveness, it is essential to provide differential protection measures by identifying an organization's information assets and classifying security levels according to the importance of the assets. The organization's document assets include various information such as core technologies, trade secrets, R&D information, and non-information. We propose a content-based information security classification method using an information network based on the information contained as well as the external attributes of document assets. The project is expected to support practical information security classification that reflects content rather than formal information security classification according to external attributes, and to improve the organization's information asset identification and classification process more efficiently through automated methods.

A Temporal and Contextual Deep Learning Intrusion Detection System for Securing Intra-vehicle Networks

Samar Bayan, Mohammad Ali Mokhadder, Utayba Mohammad, Mohamed Nafea
University of Detroit Mercy, Michigan, USA

Abstract: Modern vehicles employ intra-vehicle communication networks to exchange data between Electronic Control Units (ECUs). The most predominant protocol in this domain is the Controller Area Network (CAN) protocol. CAN networks can be accessed via multiple interfaces in a vehicle, which makes them more vulnerable to various types of cyberattacks. In this paper, a Temporal and Contextual deep learning-based Intrusion Detection System (TC-IDS) is proposed. This system uses exchanged intra-vehicle CAN messages to extract input features to be used for training, namely time intervals, data bytes and ID. The proposed algorithm incorporates two stages: training, and detection. Training is carried out offline-- using a neural network (NN) model-- to reduce computational complexity. Once the system is trained, the arriving packets are analyzed and classified as normal or hacking messages. The TC-IDS system is tested on CAN messages extracted from a 2017 production prototype Vehicle BreadBoard (VBB) and is shown to be able to detect message injection and Denial of Service attacks with high accuracy.

Determinants of Trust in Information Security of Artificial Intelligence Assistant

Hyeon Jo
Department of Strategic Planning, RealSecu, South Korea

Abstract: This research proposes the theoretical model to identify the factors influencing trust in information security of artificial intelligence assistants. Data was gathered from 232 users who had been using AIAs. This work analyzed the data by partial least square structural equation modelling (PLS-SEM) to verify the research model. The findings revealed that system quality has a significant association with trust. The results unveiled that privacy concern on surroundings is a significant enabler of trust. The findings of the present study validated that utilitarian motivation and hedonic motivation are the major antecedents of trust. Theoretical and practical implications are stated.

Drug Recommendation and Tracking using Blockchain and Machine Learning

Gokay Saldamli, Amitha S. Papetla, Amol Sadasivan, Anjali Deshmukh, Sreeja Madanambeti, Levent Ertaul
San Jose State University, San Jose, California, USA;
California State University, East Bay, Hayward, California, USA

Abstract: Pharmaceutical drug distribution is a complex process, prescription drugs are manufactured in massive volumes and then distributed to patients. One of the significant challenges that pharmaceutical industry is encountering across the globe is counterfeit drugs. In unauthorized drug distribution, the medication can be either ineffective or fatal. We are proposing a block chain-based solution to counter drug tampering. The trackability of drugs using block-chain from manufacturer to consumer ensures transparency. The block-chain-based tracking approach ensures the transactions are stored securely using blocks inside a digital ledger to ensure security. We have proposed a solution to track the delivery of drugs at each phase of the process during the whole transaction period using a distributed ledger. Furthermore, we are incorporating recommendations of drugs to a medical professional based on reviews. In the first phase, we have setup a hype-ledger fabric-based network for recording transactional data between each of the defined participants in the architecture. Hyper-ledger Fabric framework provides channels that consist of decentralized ledgers. This framework is given permissions and each participant's access to the ledger is controlled using a smart contract. We have defined all the functionalities of each of the participants in the supply chain process in the smart contracts. In the next phase, we have applied sentiment analysis on review text using tools Text-blob and Vader. We will input these sentiments to the machine learning algorithms and evaluate the algorithms; the best performing algorithm will accomplish the recommendation of medicines to a medical practitioner.

Enhancing Security Awareness on Intelligent Automated Systems for Agile Practices in Service Organisations

Olatorera C. Williams, Funminiyi Olajide, Ahmad Lotfi
Nottingham Trent University Nottingham, United Kingdom (UK)

Abstract: N/A

A Comparative Study of the Various Key Bit Sizes on RSA Algorithm with CUDA

Kwame Assa-Agyei, Funminiyi Olajide, Temitope Alade, Ahmad Lotfi
Nottingham Trent University Nottingham, United Kingdom (UK)

Abstract: N/A

Taxonomy of Bots from a Cybersecurity Perspective

Innocent Mbona, Jan Eloff
Department of Computer Science, University of Pretoria, South Africa

Abstract: Over the last couple of decades, computer programs, also known as bots, have dominated many aspects of everyday life including health, finance, entertainment, and technology amongst many others. Their dominant presence has motivated many researchers to study bots. The current literature on bots lacks a holistic view of cybersecurity threats posed by different types of bots and detection techniques. This study aims to address this limitation by performing a systematic literature review (SLR) of bots, focusing on articles published from January 2010 to December 2020. We adopted the preferred reporting items for systematic reviews and meta-analysis (PRISMA) framework to conduct the SLR. In this study, we present a high-level overview of bot cybersecurity threats by proposing a bot taxonomy based on bot type, bot cybercrime, domain type, and detection methods. This is followed by a detailed discussion of research gaps and possible future directions.

Physical Memory Attacks and a Memory Safe Management System for Memory Defense

Alon Hillel-Tuch, Aspen Olmsted
New York University Tandon School of Engineering, Brooklyn, New York, USA;
Simmons University, Boston, MA, USA

Abstract: Programming errors, defective hardware components (such as hard disk spindle defects), and environmental hazards can lead to invalid memory operations. In addition, less predictable forms of environmental stress such as radiation, thermal influence, and energy fluctuations can induce hardware faults. Sometimes, a soft error can occur instead of a complete failure, such as a bit-flip. The "natural" factors that can cause bit-flips are replicable through targeted attacks that result in significant compromises, including full privileged system access. Existing physical defense solutions have consistently been circumvented shortly after deployment. We will explore the concept of a novel software-based low-level layer that can protect vulnerable memory targeted by physical attack vectors related to bit-flip vulnerabilities.

Cluster-based Sampling Algorithm for Lightweight IoT Intrusion Detection System

Suzan Hajj, Rayane El Sibai, Adam Barada, Jacques Bou Abdo, Jacques Demerjian,
Christophe Guyeux, Abdallah Makhoul, Dominique Ginhac
ImViA, Universite de Bourgogne Franche-Comte, Dijon, France;
Faculty of Sciences, Al Maaref University, Beirut, Lebanon;
LaRRIS, Faculty of Sciences, Lebanese University, Fanar, Lebanon;
College of Business and Technology, University of Nebraska at Kearney, Kearney, Nebraska, USA;
FEMTO-ST Institute, Universite de Bourgogne Franche-Comte, CNRS, Besancon, France;
FEMTO-ST Institute, Universite de Bourgogne Franche-Comte, CNRS, Montbeliard, France

Abstract: N/A

Comparison of Tamarin-prover and ProVerif Security Verification

Takehiko Mieno, Hiroyuki Okazaki, Kenichi Arai, Yuichi Futa
Shinshu University, Japan;
Nagasaki University, Japan;
Tokyo University of Technology, Japan

Abstract: In this study, we consider incorporating an automatic verification tool into the design process for security protocols that are currently being developed. We are considering incorporating an automatic verification tool into the design process. Therefore, we employed Tamarin-prover and ProVerif, which are automatic protocol verification tools, and compared secrecy verifications compared between symmetric and public key cryptographies. From the investigation of tool features, code readability, log output of attack derivation, etc, we discussed the use of automated verification tools according to their roles.

Leveraging Software Defined Perimeter (SDP), Software Defined Networking (SDN), and Virtualization to Build a Zero Trust Testbed with Limited Resources

Dwight Horne
AT&T Center for Virtualization, Southern Methodist University, Dallas, Texas, USA

Abstract: Zero trust networks, zero trust protocol design, and zero trust software engineering are all active areas of research. Zero trust security also continues to proliferate in industry with many companies involved in one or more zero trust related projects and an Executive Order in the United States of America even mandating zero trust security with a near-term timeline for federal entities. The establishment of a zero trust testbed is a fundamental enabler to efficiently support both academic research and industry projects in the domain. In this paper, we describe how the zero trust features of a software defined perimeter (SDP) can be combined with the power and flexibility of software defined networking (SDN) and virtualization to build a zero trust testbed with limited resources. Even a student equipped with only a laptop can get started with zero trust experimentation today! We also outline useful tools for enhanced zero trust testbeds with additional (but still limited) resources and clearly show how SDP both aligns with key elements of zero trust architecture (ZTA) and contributes to satisfaction of core principles of Zero Trust by Design (ZTBD). This work will lead to additional enablers at ZeroTrustByDesign.com to facilitate design and implementation of zero trust testbeds in support of the continued evolution of zero trust research and practice. We further invite the community to join us in adding value to this growing body of zero trust knowledge and resources.

Right to Digital Privacy: An International Perspective

Esther Nehme1, Moufid El Khoury
Notre Dame University, Faculty of Law and Political Science, Zouk Mosbeh, Lebanon;
Sultan Qaboos University, College of Law, Muscat, Oman

Abstract: The right to digital privacy is established as a fundamental right in various legal instruments. Different conceptions of privacy lead digital privacy to be treated and protected differently than the fundamental right to privacy. This paper sheds light on various legal frameworks and highlights the various challenges, the enforcement of such right might encounter in practice. Similar to the commodification of data, the right to digital privacy is becoming a commodity which unequally protects data subjects. Conflicting models of data governance significantly undermine the nature and the enforcement of a digital right to privacy.

Fourth Workshop on Cybersecurity Applications and Intelligent Transportation Systems (CAITS 2022)

Chair: Prof. Pino Caballero-Gil
Department of Computer Engineering and Systems, Higher Polytechnic School of Engineering and Technology,
University of La Laguna, Spain

QuantumSolver: A Quantum Tool-set for Developers

Jose Daniel Escanez-Exposito, Pino Caballero-Gil, Francisco Martin-Fernandez
University of La Laguna, Tenerife, Spain;
IBM Research, New York, USA

Abstract: This paper introduces a new opensource quantum tool-set called QuantumSolver based on Qiskit to help developers without knowledge in quantum computing. The developed library includes a set of algorithms with different features: random number generation, Bernstein-Vazirani algorithm and quantum key distribution using the BB84 protocol. This paper described the main details about the implementation of the *toolset*, focusing in the challenges that the authors faced. Finally, this document analyzes the results obtained with some conclusions that authors compares with the included features.

Analysis and Implementation of the SNOW V

Jezabel Molina-Gil, Yanira Gonzalez-Gonzalez, Oscar Cigala-Avarez
University of La Laguna, Tenerife, Spain

Abstract: This work presents a practical analysis of the SNOW V generator. The main objective is to determine the possibility of improvement in its software implementation. It provides analyzed a software performance, to study the possibility of accelerate its implementation. Hence, a deep research is performance and it is obtained the functions' time and cpu consumption in order to determine the most expensive ones.

Study and Security Analysis of the Spanish Identity Card

Javier Correa-Marichal, Pino Caballero-Gil, Carlos Rosa-Remedios, Rames Sarwat-Shaker
University of La Laguna, Tenerife, Spain;
Government of the Canary Islands, Tenerife, Spain;
Telefonica Tech, Madrid, Spain

Abstract: The National Identity Document is a fundamental piece of documentation for the identification of citizens throughout the world. That is precisely the case of the DNI (Documento Nacional de Identidad) of Spain. Its importance has been enhanced in recent years with the addition of a chip for the authentication of users within telematic administrative services. Thus, the document has since been called: electronic DNI or simply DNIe. Sensitive user information is stored in that integrated circuit, such as personal and biometric data, along with signature and authentication certificates. Some of the functionalities of the DNIe in its current version at the time of writing this work have been implemented for years in the DNI 3.0 version launched in 2015, and therefore have already been extensively studied. This work provides a theoretical and practical compilation study of some of the security mechanisms included in the current DNIe and in some of the applications that require its use. It has been carried out using only mobile devices and generic card readers, without having any type of privileged access to hardware, software or specific documentation for the interception of packets between the DNIe and the destination application. In other words, it is an exploratory analysis carried out with the intention of confirming with basic tools the level of robustness of this very important security token.

Blockchain-based Voting System with Homomorphic Encryption

Candido Caballero-Gil, Nestor Alvarez-Diaz, Moti Yung, Pino Caballero-Gil
University of La Laguna, Tenerife, Spain; Expero, Tenerife, Spain

Abstract: Objective of this work is to design and develop a totally secure blockchain system with a web for electronic voting, but using homomorphic encryption to encrypt the vote and ensure its confidentiality and integrity in all counting and identification operations throughout the electoral process. Specifically, the application must have the ability of creating elections, registering candidates, voters and administrators, and performing the counting and custody of the votes in a totally decentralized and secure way, guaranteeing the anonymity of the votes. It seeks to find the acceptance of the authorities, since they imply a commitment with the transparency, authenticity and integrity of the entire process. In this way, the technology must overcome the challenges of the electronic voting process and reduce the doubts and suspicions of fraud that elections in general can receive. The blockchain will allow to follow the count directly, and track the destination of each vote.

Driver Facial Recognition and Drowsiness Detection

Alba Cruz-Torres, Pino Caballero-Gil, Candelaria Hernandez-Goya
University of La Laguna, Tenerife, Spain

Abstract: Traffic accidents cause constant deaths all over the world. As new applicable technology is developed, and sanctions are tightening, the number of fatalities is progressively reduced, but even so there are still too many deaths on the roads. For this reason, in this work an investigation of some of the existing solutions in the market has been carried out to increase traffic safety on roads through facial recognition of drivers and the detection of signs of drowsiness. From this analysis, the optimal design to meet the established objective has been studied. This document presents a study of Machine Learning techniques and face detection tools applied to a program for facial recognition and detection of drowsiness in drivers.

Development of a Meteorological Station and Orientation System for an IPS: Application in the ITS Field

David Suarez-Suarez, Gabriele Salvatore de-Blasio, Jose Carlos Rodriguez,
Carmelo Ruben-Garcia, Alexis Quesada-Arencibia
University of Las Palmas de Gran Canaria, Spain

Abstract: This work is focused on the field of Indoor Positioning Systems (IPS) using Bluetooth Low Energy (BLE) and WiFi technologies, proposing several modules for a robotic system that facilitates data acquisition for the development of IPS applicable to Intelligent Transport Systems (ITS). These systems can be used in public transport stations, with the aim of providing users, tourists and the general population with access to the appropriate transport to reach a specific destination. One of the research lines in this field requires a procedure that makes an exhaustive scan of the environment by means of a robotic platform collecting information on the Received Signal Strength In-dicator (RSSI) from BLE beacons and/or WiFi access points. This paper will focus on the analysis, design and implementation of a meteorological station and an orientation system to be incorpo-rated into a robotic platform. The importance of these two sub-systems is due to the fact that both meteorological variables and orientation play a key role in the nature of the signals received in IPS systems.

McEliece Cryptosystem based on Goppa Codes and LDPC Codes

Paula Perez-Pacheco, Irene Marquez-Corbella, Pino Caballero-Gil, Ignacio Garcia-Marco
University of La Laguna, Tenerife, Spain

Abstract: Post-quantum cryptography is a growing area since Shor showed (theoretically) that the quantum computer can break the most widely used cryptographic protocols today, such as RSA. Thus arises the need to study cryptosystems that can be robust against quantum computer attacks. One of them is the code-based McEliece cryptosystem created in 1978. An improved version of this cryptosystem (classic McEliece)is in the final phase of the competition created by the National Institute of Technology to update the standards and include post quantum cryptography in signature digital and encryption and key exchange. Both the original McEliece and the classic McEliece are based on Goppa codes. This study will focus, however, on the study of a cryptosystem based on the original McEliece combining Goppa codes and LDPC codes.

Drowsiness Detection in Drivers with a Smartwatch

Sonia Diaz-Santos, Pino Caballero-Gil
University of La Laguna, Tenerife, Spain

Abstract: The main objective of this work is to detect early if a driver shows symptoms of sleepiness that indicate that he/she is falling asleep and, in that case, generate an alert to wake him/her up. To solve this problem, an application has been designed that collects various parameters, through a smartwatch while driving. First, the application detects the driving action. Then, it collects information about the most significant physiological variables of a person while driving. On the other hand, given the high level

of sensitivity of the data managed in the designed application, in this work special attention has been paid to the security of the implementation. The proposed solution improves road safety, reducing the number of accidents caused by drowsiness while driving.

Preliminary Study on Post-quantum TLS

Carlos Barreda-Falciano, Pino Caballero-Gil, Jezabel Molina-Gil
University of La Laguna, Tenerife, Spain

Abstract: N/A

Cybersecurity in Voice Virtual Assistants

Eduardo Nacimiento-Garcia, Carina S. Gonzalez-Gonzalez, Candido Caballero-Gil, Francisco Gutierrez-Vela
University of La Laguna, Tenerife, Spain; University of Granada, Granada, Spain

Abstract: There are different types of Deepfakes, these can be computer-generated images, videos or audios that try to pass themselves off as real. They are made by using artificial intelligence (AI) to learn the human appearance and voice. Deepfakes can be used to create false content or to impersonate a person. In this article, we check the operation of these spoofing techniques to try to impersonate another person before Alexa and verify that unauthorized malicious activities could be carried out. We use Coqui YourTTS to clone another person's voice using a Telegram bot, which will allow us to obtain audio that tricks Alexa and her use of voice profiles to identify people.

Cybersecurity Analysis for a Smart-Home Lock

Candido Caballero-Gil
University of La Laguna, Tenerife, Spain

Abstract: As the world progresses, more and more devices are becoming smart. This means that they are able to connect to the internet and often times, be controlled remotely. While this is incredibly convenient, it also poses a new security risk that must be considered. Smart-locks are one example of a device that has become popular in recent years, but also raises security concerns. On the one hand, smart-locks can be very interesting in different situations such as when you want let somebody into your home without having to be there yourself and you can check to see if your doors are locked from anywhere in the world. However, this also means that if someone were to gain access to your account, they could easily unlock your doors and enter your home. This is a serious security risk that must be considered before using a smart-lock. Their use has led to the use of other theft techniques, such as cyber-attacks. This means that these smart-locks can be affected by public distrust. The aim of this work has been the research of security in wireless devices, focusing on the security of the Sherlock S2 Smart-lock that works via Bluetooth. Therefore, the advantages and disadvantages of having a smart-lock and its security model has been studied.

The 20th International Conference on Software Engineering Research & Practice (SERP'22: July 25-28, 2022, USA) https://american-cse.org/csce2022/conferences-SERP https://american-cse.org/csce2022/

Optimized Real-Time Assembly in a RISC Simulator

Marwan Shaban, Adam J. Rocke
Seminole State College of Florida, Sanford, Florida, USA

Abstract: Simulators for the RISC-V instruction set architecture (ISA) are useful for teaching assembly language and modern CPU architecture concepts. The Assembly/Simulation Platform for Illustration of RISC-V in Education (ASPIRE) is an integrated RISC-V assembler and simulator used to illustrate these concepts and evaluate algorithms to generate machine language code. In this article, ASPIRE is introduced, selected features of the simulator that interactively explain the RISC-V ISA as teaching aides are presented, then two assembly algorithms are evaluated. Both assembly algorithms run in real time as code is being edited in the simulator. The optimized algorithm performs incremental assembly limited to only the portion of the program that is changed. Both algorithms are then evaluated based on overall run-time performance.

Towards a Roadmap for Trustworthy Dynamic Systems-of-Systems

Rasmus Adler, Frank Elberzhager, Julien Siebert
Fraunhofer Platz 1, Kaiserslautern, Germany

Abstract: This paper gives insights into the DynaSoS project, which aims to propose a cross-domain roadmap for systems engineering research into trust-worthy dynamic systems-of-systems with autonomous components. The project follows a methodology that combines interviews and workshops with experts from various domains and a literature review. In this paper, we motivate the project, discuss five application areas with their drivers and visions, and draw initial conclusions with respect to challenges that a research roadmap should address. Our goal is to share initial results with the research community about the DynaSoS project and invite them to contribute to the derivation of the re-search roadmap.

Unified Graph Representation of Traditional and Model-driven Avionic Software Systems

Nathalie Skotnik, Andreas Frey
Department of Computer Science, Technische Hochschule Ingolstadt, Ingolstadt, Germany

Abstract: Software development for complex avionic systems is moving to a model-driven approach, where the models provide a graphical overview. However, part of the software is still developed in the traditional way or will be a reuse of already existing software. This leads to a hybrid software system, as seen from the development point of view. As the development tools are completely different, but the required quality of the software development process demands a view on the full system, this paper describes a unified graph representation that combines the artefacts from model-driven development and from traditional development. With the approach described in this paper the structure of the avionic software system can be analyzed independently from the development methodology.

Towards Systematic Software Design

Devon M. Simmonds, Emanuel S. Grant
University of North Carolina Wilmington, Wilmington, North Carolina, USA;
University of North Dakota, Grand Forks, North Dakota, USA

Abstract: Design is acknowledged as one of the more difficult activity in software engineering and academic institutions have been encouraged to put software design at the center of their computer science curricula. Systematic software design methods are needed to sup-port the successful teaching of software design in academic settings and to enhance the maturity of software engineering as an engineering discipline. This paper presents a systematic structural software design algorithm and illustrates the algorithm through a case study in software design. Case study details, results and lessons learned are presented.

Test Minimization for Mobile App Testing with FSMApp

Zeinab Abdalla, Anneliese Andrews, Ahmed Alhaddad
Computer Science Department, University of Denver, Denver, Colorado, USA

Abstract: As software Apps change, test suites are reused and updated frequently. As a result, some test cases in the test suite may become redundant when requirements covered by them are also covered by other test cases. This paper presents a test case minimization of FSMApp method for mobile Apps. It is based on concept analysis that removes redundant test cases.

Achieving Memory Safety for Unsafe Languages via Different Techniques

Mohammed Mahmoud
Department of Computer Science & Information Systems, Minnesota State University, Moorhead, Minnesota, USA

Abstract: Memory management is a common point of failure allowing memory safety violations to occur. Lower level languages such as C are considered unsafe languages, in part because of the memory safety issues that can arise from their use. Buffer overflow and dangling pointers can be caused by a simple mistake in the code. Many programming languages mitigate these errors by using garbage collection and runtime checks, but these fixes are not suitable for all applications. A variety of techniques have been employed for bringing memory safety to low level languages with minimal runtime overhead. This includes creating entirely new languages with memory safety as a first-class goal, modifying existing language syntax or semantics, and developing language-agnostic tools to verify program correctness. In this paper, I will compare the effectiveness and the limitations of each of these techniques and present a summary of the results, as well as discuss future work that can be done to improve memory safety.

A Security-privacy Based Model for Operating Systems

Mohammed Mahmoud
Department of Computer Science & Information Systems, Minnesota State University, Moorhead, Minnesota, USA

Abstract: With good cause, there has been significant effort over the past decade by operating system developers focused on improving the security of operating systems. This has resulted in regular patching, various security-oriented operating system versions being spawned, and the re-architecture of various operating system components. Recently, significant concerns about individual and institutional privacy have risen in both the media and academic arenas. There is a strong interrelationship between security and privacy even though they are distinct domains, and the aspects of each are often confused. In this paper, I will study the differences between security and privacy as it relates to operating systems and report on developments, contrasts and future improvements. I will propose a model used to collaboratively implement security and privacy concurrently within an operating system. I will discuss the results of my review and propose ways to further integrate security and privacy within operating systems.

AR-Extractor: Automatically Extracting Constraints from Android Documentation using NLP Techniques

Preethi Santhanam, Padmapriya Sakthivel, Zhiyong Shan
School of Computing, Wichita State University, Wichita, Kansas, USA

Abstract: When developing Android apps, it is difficult for the programmers to follow all programming constrains described in Android documents. This paper proposes a novel method called AR-Extractor (Android Rules Extractor) to automatically extract programming constraints from Android developer documents using natural language processing techniques. It can help programmers to reduce bugs, improve software maintainability and reliability.

A Cognitive Model for Supporting Software Functional Sizing Automation from Requirements Written as Triplets

Bruel Gerancon, Roger Nkambou, Sylvie Trudel, Serge Robert
Department of Computer Science, Universite du Quebec a Montreal (UQAM), Montreal, Canada

Abstract: The software engineering industry recognizes that cost and productivity are two key factors in the success of software projects. One of the ways that project managers can use to make a good estimation of the cost, effort, and productivity of a software project is to functionally quantify the software, that is, they measure the software functional size. Methods for measuring the functional size of software are difficult to apply in industry. Therefore, the software functional size automation from software specification documents is a priority for researchers, managers and professionals that work in the field of software production. Indeed, in the field of software development, the client generally uses his episodic memory to express his needs, with superfluous and anecdotal details, but experienced by the latter. Then, business analysts translate the needs or stories told by the customer from the episodic memory into non-formalized semantic information or requirements written in natural language, which are described in the form of Use Cases or User Stories. Thus, these software requirements are usually incomplete, inconsistent, prone to ambiguities, and difficult to be exploited by machines or by automation tools. This article introduces a cognitive model, based on first-order predicate logic, for supporting the software functional size automation from requirements written as triplets. In other words, we transformed the contents of the non-formalized semantic memory into formalized semantic requirements or formal ontology, by using description logic, in order to facilitate the software functional size automation. The formal ontology constitutes the formalized semantic memory, with automatic procedures. Subsequently, we developed a new tool (with two modules) that integrates a set of techniques in natural language processing (NLP) which helps to measure the software functional size. The first module of the tool allows to generate automatically triplets from requirements written in natural language, and transforms the data in a procedure. The second module is the functional size automation tool. The semantic and procedural memories are interconnected, which allows the automation tool to dig in these two memories to generate inferences, i.e., to determine the functional size, and identifies the types of data movements (Entry, eXit, Read, Write), as defined by the COSMIC method.Our cognitive model is an original model that combines three types of memories (episodic, semantic and procedural memories) in a complex process aimed at automating the software functional size measurement process.

Framework for the Discipline of Software Engineering

Jones Yeboah, Izunna Okpala, Sylvia Azumah, Victor Adewopo
University of Cincinnati, Cincinnati, USA

Abstract: This paper represents preliminary work in identifying the foundation for the discipline of Software Engineering and discovering the links between the domains of Software Engineering and Information Technology (IT). Our research utilized IEEE Transactions on Software Engineering (IEEE-TSE), ACM Transactions on Software Engineering and Methodology (ACM-TOSEM), Automated Software Engineering (ASE), the International Conference on Software Engineering (ICSE), and other related journal publication in the software engineering domain to address our research questions. We explored existing frameworks and described the need for software engineering as an academic discipline. We went further to clarify the distinction difference between Software Engineering and Computer Science. Through this effort we contribute to an understanding of how evidence from IT research can be used to improve Software Engineering as a discipline.

A Requirement Tracking Method Based on Information Retrieval and Source Code Structure

Tangyu Wang, Yun Wang, Zehan Tan, Weidong Yang, Hao Zhu, Shaobin Li
College of Computer Science and Technology, Fudan University, P. R. China;
Data Science Research Institute, City University of Macau, P. R. China;
Shanghai Fujia Information and Science Co. Ltd., P. R. China

Abstract: Focusing on the problems of insufficient automation, low quality of results, and poor generality in the current research on tracking methods from requirements to source code, this study proposed a technique of requirement tracking that com-bines information retrieval and source code structure information, r2cTracing-Structure, which was validated on the generic dataset. The results showed that r2cTracing-Structure can improve recall compared to the currently recognized best-performing methods while maintaining consistent precision.

Agile Effectiveness over Distributed Teams: A look into the aftermath of COVID-19

Joe R. O'Bryant, Jr., Elizabeth Schmal, Soma Datta
University of Houston - Clear Lake, Texas, USA

Abstract: Due to the COVID-19 pandemic, employers were forced to adapt to alternate work methods. Agile teams were being moved to remote, distributed telework. As the world returns to normal, the question remains of if telework is a valid work option. Telework offers flexibility for the employee and cuts costs for the employer, but is it as productive? This paper analyzes literature and COVID-19 experiences to review how effective distributed agile development is compared to pre-pandemic experiences. Then it reviews two case studies on work productivity after the-the-fact of COVID-19. At the end with the right tools in place, remote teams are as effective as the employer, management and employees make it.