CSCI 2021 BOOK of ABSTRACTS

The 2021 International Conference on Computational
Science and Computational Intelligence (CSCI'21)

https://www.american-cse.org/csci2021/
December 15-17, 2021
Luxor Hotel (MGM Property), 3900 Las Vegas Blvd. South, Las Vegas, 89109, USA

Table of Contents

Message from Program and General Co-Chairs

It gives us great pleasure to introduce this collection of papers to be presented at The 2021 International Conference on Computational Science and Computational Intelligence (CSCI'21), December 15-17, 2021, Las Vegas, Nevada, USA.

The CSCI'21 International Conference includes papers from diverse communities, including researchers from: universities, corporations, and government agencies. Accepted papers are published by Conference Publishing Services (CPS). The full proceedings/book will be published soon after the conference (like prior years). Papers published in the proceedings present solutions to problems in many important areas of computational science and computational intelligence.

Computational Science (CS) is the study of addressing problems that are impossible to solve (or difficult to solve) without computers. CS can be considered to be the bridge between computer science and other sciences. The field is interdisciplinary by nature and includes the use of advanced computing capabilities to understand and solve complex problems. In short, CS is the science of using computers to do science. Computational Intelligence (CI) is the study of computational methods in ways that exhibit intelligence. These methods adapt to changing environments and changing goals. There is a significant overlap between the fields of CI and Artificial Intelligence (AI). However, there is also a difference: in general AI techniques often involve top-to-bottom methods (i.e., methods to the solutions are imposed from the top) whereas CI techniques often involve bottom-up methods (i.e., solutions emerge from unstructured beginnings). An important part of CI includes a set of Nature-inspired computational approaches to address complex problems to which traditional methods are infeasible. Computational Science and Computational Intelligence, both share the same objective: finding solutions to difficult problems. However, as stated earlier, the methods to the solutions are different.

Considering the above broad outline, the International Conference on Computational Science and Computational Intelligence (CSCI'21) is composed of the following topical symposiums: Computational Science (CSCI-ISCS); Computational Intelligence (CSCI-ISCI); Computational Biology (CSCI-ISCB); Cyber Warfare, Cyber Defense, & Cyber Security (CSCI-ISCW); Artificial Intelligence (CSCI-ISAI); Smart Cities and Smart Mobility (CSCI-ISSC); Big Data and Data Science (CSCI-ISBD); Education - CS & CE (CSCI-ISED); Health Informatics and Medical Systems (CSCI-ISHI); Mobile Computing, Wireless Networks, & Security (CSCI-ISMC); Software Engineering (CSCI-ISSE); Internet of Things & Internet of Everything (CSCI-ISOT); Social Network Analysis, Social Media, & Mining (CSCI-ISNA); Cloud Computing and Data Centers (CSCI-ISCC); Parallel & Distributed Computing (CSCI-ISPD); and Signal & Image Processing, Computer Vision & Pattern Recognition (CSCI-ISPC).

The main objective of the CSCI Conference is to facilitate increased opportunities for cross-fertilization across CS and CI. The CSCI Conference is committed to encouraging diversity and eliminating discrimination in both its role as a conference and as a provider of services. CSCI aims to create a culture that respects and values each others' differences, that promotes dignity, equality and diversity, and that encourages individuals to develop and maximize their true potential. We are committed wherever practicable to organizing a conference that broadly reflects the international community. We hope that we have achieved these important objectives.

The Steering Committee and the Program Committee would like to thank all those who submitted papers for consideration. The conference had paper submissions from 62 countries. About 57% of the submissions were from outside the United States. Each submitted paper was peer-reviewed by at least two experts in the field for originality, significance, clarity, impact, and soundness. In cases of contradictory recommendations, a member of the conference program committee was charged to make the final decision; often, this involved seeking help from additional referees. In addition, papers whose authors included a member of the conference program committee were evaluated using the double-blinded review process. One exception to the above evaluation process was for papers that were submitted directly to chairs/organizers of sessions/workshops; in these cases, the chairs/organizers were responsible for the evaluation of such submissions. The overall paper acceptance rate for regular and short papers was 17%; and 18% of the remaining papers were accepted as extended abstract (poster) papers (at the time of this writing, we had not yet received the acceptance rate for two research tracks.)

We are very grateful to the many colleagues who offered their services in organizing the conference. In particular, we would like to thank the members of the Program Committee and the Steering Committee of CSCI'21. The members of the committees will be requested (after the conference) to provide their expertise and services for selecting papers for publication (extended versions) in

various research book series (to be prepared for publishers including: Springer, Elsevier, and others). We would also like to thank the main sponsor of the conference: American Council on Science & Education.

We express our gratitude to all speakers and authors - the list of speakers appears in the conference schedules. We would also like to thank the followings: the staff of the Luxor hotel (Conference division); and Conference Publishing Services of IEEE Computer Society production editors and managers: Javier Gurrola, Patrick Kellenberger, Lisa O'Conner, and Lorretta Palagi.

We present the proceedings of CSCI'21.

Professor Hamid R. Arabnia - Editor-in-Chief, The Journal of Supercomputing (Springer); Fellow & Adviser of Center of Excellence in Terrorism, Resilience, Intelligence & Organized Crime Research (CENTRIC); Professor Emeritus of Computer Science, The University of Georgia, Georgia, USA

Professor Leonidas Deligiannidis - Professor, School of Computing and Data Science, Wentworth Institute of Technology, Boston, Massachusetts, USA

Professor Fernando G. Tinetti - Facultad de Informatica, Universidad Nacional de La Plata, La Plata, Argentina; Researcher CIC Prov. Bs. As., Argentina; Recipient of the NSF-TCPP Early Adopter Award, NSF/IEEE-TCPP Curriculum Initiative on Parallel and Distributed Computing

Professor Quoc-Nam Tran – Department of Computer Science, Southeastern Louisiana University, Louisiana, USA

Co-Chairs, Steering Committee, CSCI'21

Conference Organization – CSCI 2021

Steering Committee Co-chairs

Professor Hamid R. Arabnia - Editor-in-Chief, The Journal of Supercomputing (Springer); Fellow & Adviser of Center of Excellence in Terrorism, Resilience, Intelligence & Organized Crime Research (CENTRIC); Professor Emeritus, Department of Computer Science, The University of Georgia, Georgia, USA

Professor Leonidas Deligiannidis - Professor, School of Computing and Data Science, Wentworth Institute of Technology, Boston, Massachusetts, USA

Professor Fernando G. Tinetti - Facultad de Informatica, Universidad Nacional de La Plata, La Plata, Argentina; Researcher CIC Prov. Bs. As., Argentina; Recipient of the NSF-TCPP Early Adopter Award, NSF/IEEE-TCPP Curriculum Initiative on Parallel and Distributed Computing

Professor Quoc-Nam Tran – Department of Computer Science, Southeastern Louisiana University, Louisiana, USA

Members of Steering Committee

Professor Babak Akhgar - Fellow of the British Computer Society, CITP, Professor of Informatics, Co-Director of CENTRIC (Centre of Excellence in Terrorism, Resilience, Intelligence & organised Crime research), Sheffield Hallam University, Sheffield, United Kingdom

Professor Abbas M. Al-Bakry - University President, University of IT & Communications, Baghdad, Iraq

Professor Emeritus Nizar Al-Holou - Electrical and Computer Engineering; Vice Chair, IEEE/SEM-Computer Chapter, University of Detroit Mercy, Michigan, USA

Professor Rajab Challoo - Professor and Chairman, EECS Department; Director, Center for Controls, Robotics, Intelligent Systems and Sensors Research, CRISS; President, Faculty Senate; Former Chair of Computer Society (for 7 years); Texas A&M University-Kingsville, Kingsville, Texas, USA

Professor Chien-Fu Cheng - Professor of CS, Department of Computer Science and Information Engineering, Graduate Institute of Networking & Communication, Tamkang University, Taiwan, R.O.C.

Professor Hyunseung Choo - Director of Intelligent HCI Convergence Research Center (ITRC) of Ministry of Information and Communication; International Advisory Chair, ACM RACS; College of Information and Communication Engineering, Professor of IS, Department of Interaction Science, Sungkyunkwan University (SKKU), Suwon, Republic of Korea

Professor Emeritus Kevin Daimi - Computer Science and Software Engineering Programs, Department of Mathematics, Computer Science and Software Engineering, University of Detroit Mercy, Detroit, Michigan, USA

Professor Eman M. El-Sheikh - Director, Center for Cybersecurity; Member of AWIS (Association for Women in Science); Professor of CS, Department of Computer Science, University of West Florida, Pensacola, Florida, USA

Professor Mary Mehrnoosh Eshaghian-Wilner – Adjunct Professor of Electrical Engineering, University of California Los Angeles, UCLA, California, USA

Dr. David L. Foster - Associate Professor, Computer Engineering; Kettering University, Michigan, USA

Members of International Program Committee (CSCI 2021)

- **Professor Afrand Agah** - Graduate Coordinator, Department of Computer Science, West Chester University of Pennsylvania, Pennsylvania, USA
- **Bharat Bhushan Agarwal** - Reader, IFTM University, Moradabad, India
- **Dr. Omaima Nazar Ahmad AL-Allaf** - CIS Deptartment, Faculty of Sciences and Information Technology, Al-Zaytoonah University of Jordan, Amman, Jordan
- **Professor Wasim A. Al-Hamdani** - Professor of Cryptography and InfoSec; Information Security Lab.; Computer Science; Kentucky State University, Frankfort, Kentucky, USA
- **Dr. Ismail Khalil Al Ani** - Dean, College of Engineering and Computer Science, Ittihad University, Ras Al Khaimah, United Arab Emirates (UAE)
- **Professor Mehran Asadi** - Information Technology; Department of Business and Entrepreneurial Studies, Lincoln University, Pennsylvania, USA
- **Dr. Dr. Travis Atkison** - Assistant Professor, Computer Science, The University of Alabama, Tuscaloosa, Alabama, USA
- **Dr. Azita Bahrami** - President, IT Consult, Georgia, USA
- **Dr. Mehdi Bahrami** - IEEE Senior Member; Fujitsu Laboratories of America, Inc., California, USA
- **Dr. Petra Saskia Bayerl** - Department of Organisation and Personnel Management, Rotterdam School of Management (RSM), Erasmus University Rotterdam, Holland, The Netherlands
- **Professor Jane M. Binner** - Chair in Finance, Department of Accounting and Finance, Birmingham Business School, University of Birmingham, Birmingham, West Midlands, United Kingdom
- **Professor Dr. Juan-Vicente Capella-Hernandez** - Universitat Politecnica de Valencia (UPV), Department of Computer Engineering (DISCA), Valencia, Spain
- **Professor Juan Jose Martinez Castillo** - Director, The Acantelys Alan Turing Nikola Tesla Research Group and GIPEB, Universidad Nacional Abierta, Venezuela
- **Dr. Rui Chang** - Lab Director; Department of Genetics and Genomic Science, Icahn Institute for Genomics and Multiscale Biology, Mount Sinai School of Medicine, New York, USA
- **Professor Dongsheng Che** - Director, Che Lab of Bioinformatics at ESU, Department of Computer Science, East Stroudsburg University of Pennsylvania, Pennsylvania, USA
- **Dr. Jianhung Chen** - Department of Computer Science and Information Engineering, Chung-Hua University, Hsin-Chu, Taiwan
- **Professor Mu-Song Chen** - Department of Electrical Engineering, Da-Yeh University, Changhua, Taiwan, R.O.C.
- **Dr. Xin (Thomas) Chen** - Research Corporation of the University of Hawaii, Institute for Astronomy, Honolulu, Hawaii, USA
- **Professor Steve C. Chiu** - Program Director of Electrical Engineering, College of Engineering, Idaho State University, Idaho, USA
- **Professor Mark Yul Chu** - Department of Electrical Engineering, The University of Texas Rio Grande Valley, Texas, USA
- **Professor Ann Leslie Claesson-Vert** (Vice Chair) - School of Nursing, Northern Arizona University, North Valley Campus, Phoenix, Arizona, USA
- **Dr. Jose Alfredo F. Costa** - Member of IEEE Computational Intelligence Society; President, SBIC - Brazilian Computational Intelligence Society; Associate Professor, Department of Electrical Engineering, UFRN - Universidade Federal do Rio Grande do Norte (Federal University - UFRN), Brazil
- **Dr. Arianna D'Ulizia** - Institute of Research on Population and Social Policies, National Research Council of Italy (IRPPS), Rome, Italy
- **Professor Zhangisina Gulnur Davletzhanovna** - Vice-rector of the Science, Central-Asian University, Kazakhstan, Almaty, Republic of Kazakhstan; Vice President of International Academy of Informatization, Kazskhstan, Almaty, Republic of Kazakhstan
- **Professor Noel De Palma** - Associate Vice President for Research, Computer Science; Leader, ERODS Research, LIG Labs.; Board of Directors, OW2 Consortium; Department of Computer Science; University of Grenoble I, Grenoble, France

Note that the title of papers and authors' names that appear in the "Book of Abstracts" were extracted from the papers that were submitted to the EVALUATION web site. The official published proceedings/book will have any and all changes/revisions that authors may have done to the title and/or authors lists in the final version of their manuscripts.

CSCI-ISOT:
INTERNET OF THINGS & INTERNET OF EVERYTHING

DDS-Cerberus: Ticketing Performance Experiments and Analysis

Andrew T. Park, Richard Dill, Douglas D. Hodson, Wayne C. Henry
Department of ECE, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, USA

Abstract - Data Distribution Service (DDS) is a publishsubscribe middleware used to distribute data between real-time systems, production environments, and small embedded platforms. In DDS, Nodes have at least one Publisher or Subscriber. Publishers and Subscribers use unique Topics to send and receive messages. Each Subscriber has permission to read the Publisher's message if it references the same Topic sent from the Publisher. This capability supports real-time communication, sacrificing security, such as impersonation attacks. This paper details, tests, and evaluates DDS-Cerberus (DDSC), a novel distributed communication protocol that integrates the Kerberos ticketing system with DDS. DDS-C integrates Kerberos authentication and Ticket retrieval with Publishers and Subscribers. Experiments have six parameters each with a 2:1 Publisher to Subscriber ratio. Performance tests modify the message byte size to emulate .txt and .mp3 files: 10 KB, 100 KB, 1 MB, 5 MB, 10 MB, and 20 MB. Experiment metrics for functionality and performance are the messages per second and latency in a wired environment. Experiments utilize ROS 2 (Robot Operating System) as a testbed. Initial tests for a baseline are conducted without DDS modifications and subsequent tests with DDS-C modifications. The results reveal that due to the ticketing component, DDS-C increases DDS security by preventing impersonation attacks while negligibly increasing average processing compared to baseline results.

A Service Tier Design for the EMULSION IoT Platform

Ivan Ganchev, Zhanlin Ji, Mairtin O'Droma
University of Plovdiv "Paisii Hilendarski", Plovdiv, Bulgaria;
Institute of Mathematics and Informatics - Bulgarian Academy of Sciences, Sofia, Bulgaria;
North China University of Science and Technology, Caofeidian, Tangshan City, China;
Telecommunications Research Centre, University of Limerick, Limerick, Ireland

Abstract - This paper presents an IoT service tier design and its setting within the multi-tiered architectural structure of EMULSION, the newly proposed horizontal-type IoT platform. Besides briefly outlining the key rationale for this IoT platform paradigm, design details are presented on the two corresponding sub-tiers, facilitating the efficient and effective recommendation and use of IoT services by consumers, considering the current context so as to avail of the 'best' service instances under the always best connected and best served (ABC&S) communication paradigm.

A Novel Modular Low Power and Low Cost IoT Wireless Sensor Node for Air Quality Monitoring

Matthew Meli, Edward Gatt, Owen Casha, Ivan Grech, Joseph Micallef
Department of Microelectronics and Nanoelectronics, University of Malta, Msida, Malta

Abstract—The design and implementation of a novel modular low power and low cost IoT wireless sensor node for air quality monitoring is presented. The modular device is made up of tow boards: a power management and microcontroller board and an air quality sensors and radio shield. The fabricated device achieves state of the art low power performance using a fully fragmented dynamic power management architecture. The device consumes 270 nA in sleep mode and an average of 38 mA in active mode. The average power consumption per hour is 327 µAh. This results in a 40-month battery autonomy using a 10,500 mAh battery.

Analyzing a Low-bit Rate Audio Codec - Codec2 - on an FPGA

Peter Jamieson, Santhiya Sampath Kumar
Electrical and Computer Engineering, Miami University, Oxford, Ohio, USA

Abstract—Audio compression codecs are an important application in the Internet of Things (IoT) space where small sensing devices may gather voice signals, but then need to transmit the information to aggregating servers at a low cost. In this work, we implement and evaluate a hardware implementation of the Codec2 (a lossy speech compression codec) in Verilog Hardware Description Language (HDL) and map it to an Intel CycloneIV FPGA. We describe the details of our implementation approach, including how we convert C code of Codec2, how we represent data inside the hardware implementation and the associated cost of this implementation on an FPGA. We analyze our implementation compared to a microprocessor implementation of the original software to observe what performance we get on an FPGA versus a microprocessor. Our hardware implementation of Codec2 is qualitatively the same in terms of hearing the spoken transmission and has an error rate of 6.55 bits per frame (48 bits), and is 10 times slower and consumes slightly more power on a small FPGA compared to a microprocessor implementation. We provide this design as open-source hardware so that we have an HDL version of this application that can be mapped to both FPGAs and ASICs providing a research reference point, a baseline to optimize, and a described methodology to convert C to hardware for researchers.

Scalability Evaluation of a Per-User Access Control Framework

Arata Endo, Chunghan Lee, Susumu Date
Cybermedia Center, Osaka University, Osaka, Japan;
Infotech, Toyota Motor Corporation, Tokyo, Japan

Abstract—Today's Internet of Things (IoT) devices have a variety of security requirements and policies. While an access control is applied to such devices to meet the varieties of requirements and policies, the access control has rarely been used for network resources. Due to this situation, we have proposed a per-user access control framework, which realizes the access control for network links and bandwidth as network resources by using Software-Defined Networking, in our previous work. The proposed framework enables a network administrator to apply access control to network resources simply by giving the administrator's policy as input to the proposed framework. However, there remains the concern that the proposed framework may cause a significant overhead for the data transfers when the number of IoT devices is increased. In this paper, we investigate how scalable the proposed framework is as infrastructure, by considering the actual and practical situation where lots of IoT devices are used. Our evaluation results imply that the overhead incurred by the proposed method is negligible, especially in the case where IoT devices transfer large-sized data. Also, the evaluation results show that the proposed framework reduces the exposure time of the IoT devices to a third party.

IoT-based Monitoring System to Prevent Collapse of Falsework Systems During Construction

R.C.Y. Lam, L.C.H. Lam, N.X. Li, P.K.K. Lee
Department of Civil Engineering, President's Office, Chu Hai College of Higher Education, Hong Kong, China

Abstract—A recurring accident that often involves human casualties during construction projects is the collapse of falsework / temporary support systems during demolition and new construction works. This paper describes an IoT monitoring system designed to prevent such failures through detecting and reporting, in real time, abnormalities that would precede an accident. Practical applications of the monitoring system demonstrate that IoT monitoring provides an effective preventive measure suitably introduced to industry practices.

Service Prototype Provisioning for the EMULSION IoT Platform

Ivan Ganchev, Zhanlin Ji, Mairtin O'Droma
University of Plovdiv "Paisii Hilendarski", Plovdiv, Bulgaria;
Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria;
College of Artificial Intelligence, North China University of Science and Technology, Caofeidian, Tangshan City, China;
Telecommunications Research Centre, University of Limerick, Limerick, Ireland

Abstract - This paper presents the multi-tiered architectural structure of the newly proposed horizontal-type IoT platform, called EMULSION. Besides briefly outlining the key rationale for this IoT platform paradigm, design aspects are presented with respect to the service prototype provisioning, facilitating the efficient and effective recommendation and use of IoT services by consumers, considering the current context so as to avail of the 'best' service instances under the always best connected and best served (ABC&S) communication paradigm.

Safe Social Internet of Thing for Urban Mobility Services

Matteo Anedda, Mauro Fadda, Massimo Farina, Roberto Girau, Mariella Sole, Daniele D. Giusto
University of Cagliari, Italy; University of Sassari, Italy; University of Bologna, Italy

Abstract—Smart cities are characterized by smart heterogeneous devices that can interact and cooperate with each other by exchanging regularly big amounts of data with the big issue to treat sensitive data in a properly respectful manner, avoiding exposure to the risks that new technologies inevitably bring to the fore. This objective can only be pursued with adequate knowledge of the risks and methods of protection, for this reason in addition to producing materially functional results, it has been studied in depth the techniques of protection of personal data including the anonymization and pseud-anonymization of sensitive data. We provide the analysis of the state of the art that starts from the concepts of security and privacy and comes to an analysis of anonymization algorithms. This analysis tries to give an overview of the two fundamental issues in the field of data security: privacy, according to the European Regulation 2016 (GDPR) and the practical techniques with which it is preserved, with particular attention to anonymization algorithms: we analyze the advantages and disadvantages of the latter. The performance of the proposed solution is compared against that of a TraffictYpe-based DifferEntiated Reputation (TYDER) algorithm. This performance was evaluated in terms of QoS parameters such as delay, latency, packet loss and prediction error. The results show how MISSION outperforms TYDER in urban mobility scenario.

Development of Wireless Communication System for Reliable Acoustic Data Collection Toward Anomaly Detection on Mechanical Equipment

Chinari Takano, Shinya Fujino, Daiki Nobayashi, Kazuya Tsukamoto, Mitsunori Mizumachi, Takeshi Ikenaga
Kyushu Institute of Technology, Fukuoka, Japan

Abstract—With the proliferation of Internet of Things (IoT) devices that can send and receive data via wireless communication in this decade, we are able to monitor and operate these devices remotely. An example of an IoT system using wireless communication is a system for anomaly detection on the mechanical equipment by the acoustic data. In order to detect anomalies using acoustic data, continuous recording is essential, thereby increasing the data size. Although the Wi-Fi network provides us with high-capacity data transfer, the performance degradation cannot be avoided due to several reasons such as packet losses caused by collisions with other devices using the same frequency and the increase in the distance between communicating 2 devices. In this paper, we develop a wireless communication system for reliable acoustic data collection toward anomaly detection on mechanical equipment. First, as a preliminary experiment, we investigate the communication characteristics for transmission of large-size data by Wi-Fi in indoor and outdoor environments. As results of the experiments, the communication performance was insufficient for transferring all recorded data handled by this system. Based on the results of our preliminary experiments, we propose a simple and heuristic transmission timing control method and the method that can reduce the amount of transmission data for realizing a stable acoustic data collection system. Finally, through demonstration experiments consisting of multiple mechanical equipment in the real field, we verified the feasibility of the acoustic data collection system.

UV-C VentGuard: An IoT-based Monitoring and Disinfection System for Pathogen Inactivation in Ventilation Pipes of High-rise Buildings

R.C.Y. Lam, C.K. Tang, K.L.A. Cheung, L.C.H. Lam
Department of Civil Engineering, President's Office, Chu Hai College of Higher Education,
Hong Kong Baptist University, Hong Kong, China;
Department of Biology, Hong Kong Baptist University, Hong Kong, China

Abstract—SARS-CoV-2 traveling through ventilating pipes in high-rise buildings present an urgent concern to address. This paper describes an IoT monitoring system designed to disinfect the air traveling through the pipes. Site tests demonstrate that the system provides a cost-effective solution for pathogen inactivation in ventilating pipes of high-rise buildings, and that it can play a positive role in mitigating the spread of the COVID-19 pandemic in built environments.

Identifying Actionable Customer Behavior through Advanced Analysis of Bank Transaction Data

Anna Nesvijevskaia, Vincent Martenot, Paul Fogel, Jean-Daniel Zucker
Laboratory DICEN IDF, CNAM, Quinten, Paris, France;
DataLab, Quinten, Paris, France;
DataLab Senior Advisor, Quinten, Paris, France;
IRD, UMMISCO, Sorbonne University, Bondy, France

Abstract—Artificial Intelligence has opened new doors for customer relationship personalization by capturing life events to tailor front and back-office interactions. Individual bank account data are particularly rich in information on these life events, but few banks have gone beyond its basic use. In this paper, we describe an innovative and original methodological framework to give meaning to bank transactions and make them actionable under operational and regulatory constraints. The approach includes unsupervised methods that limit upstream feature engineering and are based on a global modeling of a customer's journey through sequence objects.

A Token Label for Maturity Prediction using Current-Future Changing Rate

Alisdair Chun On Lee, Shum Paul Kwok-Ching, H. Anthony Chan
Computing and Information Sciences, Caritas Institute of Higher Education, Hong Kong;
Business and Hospitality Management, Caritas Institute of Higher Education, Hong Kong

Abstract—Fruit inside a container is difficult to disclose enough information to allow us to acknowledge its future expiration in real-time in the logistics and fruits transportation. Primarily owing to this, we notice the food expires after the expired signal appears, causing 40% of global fruit waste. This work proposes a maturity token label using the current-future changing rate with a sequence to sequence model to regularly adapt to each scenario of the surrounding environment during the maturity life-cycle. The intelligent container then notifies us about the future expired signal to enable remedy in transportation to prevent fruit loss.

A Conceptual Framework for a Blockchain-based Tax Payment Financial Service

Abhishek Phadke, F. Antonio Medrano, Jaimin Brahmbhatt
Department of Computer Science, Texas A&M University, Corpus Christi, Texas, USA

Abstract—Blockchain promotes secure and transparent data flow, and its distributed nature makes it ideal for conducting transactions visible to the public. Tax payments by citizens are the primary source of funds for governments across the globe. However, the system is overburdened and plagued by problems making tax recovery a lengthy and expensive process. Blockchain-based applications and tokens are long thought to be ideal for conducting financial operations. This study proposes a conceptual blockchainbased framework that aims to resolve issues in the current tax collection and recovery system deployed by governments across the globe. Blockchain has provided the public with new and innovative ways to invest and hold wealth, and the creation of an integrated taxation service based on decentralized records and reverse token payments can help modernize these outdated government financial services.

Weather Information Acquisition System using Small Wireless Microcomputers and Environmental Sensors for Agricultural Workers

Uyou Cyou, Shohei Yokoyama, Shingo Otsuka
Graduate School of Kanagawa Institute of Technology, Kanagawa, Japan;
Department of Computer Science, Tokyo Metropolitan University, Tokyo, Japan;
Department of Information and Computer Sciences, Kanagawa Institute of Technology, Kanagawa, Japan

Abstract—The empirical rules of agricultural workers have been inherited successors in Japan. However, the empirical rules have been lost by the deserted garden due to the death of the agricultural workers and the lack of successors. We construct the low-cost system for acquiring meteorological data using small wireless microcomputers and environmental sensors as a solution to this problem. In this paper, we describe a proposal system, and then describe various errors that occurred in the demonstration experiment. After that, we describe our solutions. In addition, we also show the visualization of the collected environmental data.

CSCI-ISCW:
Cyber Warfare, Cyber Defense, & Cyber Security

On the Impact of the Embedding Process on Network Resilience Quantification

Luiz Manella Pereira, S. S. Iyengar, M. Hadi Amini
Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, Florida, USA;
Sustainability, Optimization, and Learning for InterDependent Networks Laboratory (solid lab),
Florida International University, Miami, Florida, USA

Abstract— Network resilience is crucial to ensure reliable and secure operation of critical infrastructures. Although graph theoretic methods have been developed to quantify the topological resilience of networks, i.e., measuring resilience with respect to connectivity, in this study we propose to use the tools from Topological Data Analysis (TDA), Algebraic Topology, and Optimal Transport (OT). In our prior work, we used these tools to create a resilience metric that bypassed the need to embed a network onto a space. We also hypothesized that embeddings could encode different information about a network and that different embeddings could result in different outcomes when computing resilience. In this paper we attempt to test this hypothesis. We will utilize the WEGL framework to compute the embedding for the considered network and compare the results against our prior work, which did not use an embedding process. To our knowledge, this is the first attempt to study the ramifications of choosing an embedding, thus providing a novel understanding into how to choose an embedding and whether such a choice matters when quantifying resilience.

Deep-Learning Side-Channel Attack Against STM32 Implementation of AES

Fanliang Hu, Huanyu Wang, Junnian Wang
School of Physics and Electronic Science, Hunan University of Science and Technology, Xiangtan, Hunan, China;
School of EECS, KTH Royal Institute of Technology, Stockholm, Sweden

Abstract—Deep-Learning Side-Channel Attacks (DLSCAs) have become a realistic threat to cryptographic algorithms, such as Advanced Encryption Standard (AES). Since the encryption has to run in hardware at some point to actually do things, there might be some unintentional physical leakage, such as the different amount of power consumed by the victim device. By using deep-learning models to analyze the power traces, the attacker is able to derive the secret key. In this project, we implement a real deep-learning based attack to against a STM32 implementation of AES. We apply four different types of neural networks, MLP, CNN, LSTM and RNN, to classify traces. Afterwards, we evaluate to which extent different types of models could make the attack more efficient.

A Novel Traceback Technology for E-mail Sender Verification

Woo Young Park, Sang Hyun Kim, Duy Son Vu, Chang Han Song, Hyeon Jo, Hee Soo Jung
Security Laboratory, Realsecu, Busan, South Korea;
Headquarter, Realsecu, Busan, South Korea

Abstract— Since the COVID-19 outbreaks, the non-face-to-face work environment has been spread and malicious e-mail attacks have also increased. In particular, the damage caused by impersonation, forgery, and alteration of e-mail has been drastically increasing. While the existing e-mail security solutions scan and examine large mails, advertising mails, and junk mails as major targets, they are not able to identify or distinguish the sender of impersonated or forged mail. This study developed a novel e-mail processing system based on SMTP (Simple Mail Transfer Protocol) reply codes. In the proposed system, the SMTP reply codes related to the loss of e-mail are used to perform traceback and verify the sender of e-mails. When an e-mail is received, the system generates verification request information and conducts communication to the sender's e-mail server and DNS. According to information that the sender's e-mail server offers back, the proposed system discriminates the authenticity of e-mail senders. The proposed system was introduced to actual institutions and performance verification was carried out. As a result, the proposed system has significantly improved the rate of e-mail sender verification. We hope that the results of this study can serve as a useful guideline for the development of e-mail security.

A New SOCMINT Framework for Threat Intelligence Identification

Marco San Biagio, Roberto Acquaviva, Valentina Mazzonello, Ernesto La Mattina, Vito Morreale
Homeland Security research unit, Engineering Ingegneria Informatica s.p.a., Palermo, Italy

Abstract—Social Media platforms has changed the way of communication between people. These platforms allow users and organizations to communicate each other, sharing their ideas, interests, and knowledge, becoming a part of their daily routine. On the other hand, social media platforms represent good channels to exploit for perpetrating various kind of crimes and crimesillegal activities in several criminal areas, including cyber-crimes and terrorism. The goal of this work is to present a novel framework for Threat Intelligence that adopting Machine Learning and Artificial Intelligence techniques for OSINT (Open–source intelligence) investigation, is able to extract actionable intelligence for threats.

A Safe Dynamic Access Control Providing Mandatory Automotive Cybersecurity

Vincent Hugot, Adrien Jousse, Christian Toinard, Benjamin Venelle
LIFO, INSA Centre Val de Loire, Bourges, France;
GEEDS, Valeo, Creteil, France

Abstract—The recent computerization of automotive vehicles offers new ways for attackers to penetrate critical systems. To preserve their safety, cybersecurity is required. Cybersecurity can be enforced with different mechanisms, such as access control, thanks to reference monitors. In certain scenarios, systematic access control may harm safety properties. Therefore, cybersecurity of automative vehicles remains an open problem. Our paper provides a general approach to enforce security properties while guaranteeing their innocuousness wrt. the vehicle's safety. Our approach considers how to add dynamic mandatory access control in an automotive system while preserving the safety of the vehicle. This is achieved by first modeling the system. Then, a mandatory access control is proposed wrt. the different vehicle states. Using a dedicated model checking, the safety of the online control is verified regarding a formal attack model. Our approach eases the improvement of both the mandatory control and attack assumptions.

The H2020 ANITA Platform: Generating Knowledge about Crime through User-centred Innovative Tools

Valentina Mazzonello, Ernesto La Mattina, Roberto Acquaviva, Gianluca Cina, Marialuna De Tommaso, Youssef Bouali, Barbara Pirillo, Simone Simoncini
Research and Innovation Division, Engineering Ingegneria Informatica S.p.A., Palermo, Italy;
Research and Innovation Division, Engineering Ingegneria Informatica S.p.A., Napoli, Italy;
Research and Innovation Division, Engineering Ingegneria Informatica S.p.A., Lecce, Italy

Abstract - This paper illustrates a novel user-centred and secure investigation platform to discover relevant data sources disseminated on the Web (including the Dark Web) and to analyse, enrich and correlate them to support knowledge-generation and investigations on illegal trafficking activities. The platform was designed and implemented in the context of the Horizon 2020 project ANITA (https://www.anita-project.eu/) to effectively support Law Enforcement Agencies in better understanding and investigating the online illegal trafficking framework and further improve their capacity to analyse complex criminal scenarios and big amount of relevant data. The paper illustrates the ANITA platform capabilities through some of the 'horizontal criminal scenarios' defined by the Project to explore connections between illegal trafficking activities concerning different goods. These 'scenarios' refer to cases of the same vendor distributing both firearms and synthetic drugs on the same or in different crypto markets, the same profiles found to be active on both some Dark Web crypto markets and Surface platforms or terrorist attacks performed with firearms bought on the Dark Web.

Zero-Trust Model of Cybersecurity: A Significant Challenge in the Future

Fadi Al-Ayed
Department of Electrical Engineering and Computer Science, The Catholic University of America, Washington D.C., USA

Abstract — To address future cybersecurity challenges, this paper proposes a real-time three-factor authentication scheme (RT3FA). The model integrates the characteristics of multi-factor authentication and real-time actual information. As face biometrics are needed in addition to two-factor authentication, the additional layer of protection raises the obstacles for data access. Facial biometric is accomplished by synchronizing real-time information with feature recognition via an instantaneous live feed from the user's camera. However, the improved protection may cause efficiency issues and thus, require higher capacities for both the user's device and the database system.

Detecting Ransomware Automated Based on Network Behavior by Using Machine Learning

Haydar Teymourlouei, Vareva E. Harris
Department of Technology & Security, Bowie State University, Bowie, Maryland, USA;
Criminal Justice Administration and Center for Cybersecurity, Benedict College, Columbia, South Carolina, USA

Abstract - Ransomware has been a detrimental form of malware, in which many companies have become victims of these attacks and are required to release specific amounts of money to attackers without knowing if they will ever relieve their data. There has been a dramatic growth in ransomware attacks in recent years. This research will provide effective methods for preventing these attacks and limiting the effects of the attack if they ever occur. Our technique is based on using an ensemble machine learning classification algorithm in a random forest and boosting algorithm to evaluate network behavior. We used the Adaboost algorithm to create a sequence model to predict accurate results in an automated effective, and efficient method that detected ransomware.

Leveraging Information Security Continuous Monitoring to Enhance Cybersecurity

Tina AlSadhan, Joon S. Park
School of Information Studies, Syracuse University, Syracuse, New York, USA

Abstract— Cyber attacks against U.S. Federal information systems are relentless and increasingly sophisticated. The probability for grave damage continues to escalate despite the efforts and significant resources expended. The detection, analysis, and prioritization of cybersecurity vulnerabilities, threats, and the status of the effectiveness of cybersecurity protection measures is needed within minutes to reduce or eliminate compromise and the associated debilitating consequences. Information Security Continuous Monitoring (ISCM) leverages technology to evolve from compliance-focused cybersecurity to data-driven risk management. ISCM enables real-time or near-real-time cyber situational awareness to be responsive to the explosive rates of vulnerabilities, persistent threats, and determined enemies. Despite the promises of ISCM, Federal agencies continue to face challenges in achieving effective ISCM. Therefore, as part of our ongoing research, in this paper we discuss the current ISCM status, analyze its challenges for cybersecurity, and propose the future strategies to leverage ISCM for cybersecurity.

Quantifying the Impact of Vulnerabilities of the Components of an Information System Towards the Composite Rise Exposure

Yanzhen Qu, Adam English, Bruce Harmon
College of Computer Science, Engineering, and Technology, Colorado Technical University, Colorado Springs, Colorado, USA;
Ritchie School of Engineering and Computer Science, University of Denver, Denver, Colorado, USA

Abstract— To compensate the lacking of any concrete scoring formula of the CVSS v3 for the category of "Environment", in this full research paper, we present a novel formula for objectively quantifying composite vulnerability exposures for non-terminal components of an information system. The paper examines limitations of the CVSS v3 calculator definition, notably the capacity to characterize vulnerabilities from a composite perspective, providing a means to output a composite CVSS-compliant vulnerability score for aggregated system components. We provide the definitions for related concepts, formulas for determining component vulnerability, and a formula for calculating composite vulnerability. The common implementation of a Linux, Apache, MySQL, PHP (LAMP) stack provides a practical demonstration of the foundational formulas.

Taxonomy for Malware Detection to Enhance the Security of Smart Devices using AI

Anusha Thapa, Abeer Alsadoon, Sreekanth Gopalakrishnan Nair, Muhammad Siddiqi, Rashidul Mubasher,
Rajesh Ampani, Benoy Varghese, P.W.C. Prasad
Study Group Australia, Sydney, Australia;
Kent Institute Australia, Information Technology Department, Sydney, Australia;
Asia Pacific International College (APIC), Information Technology Department, Sydney, Australia;
Advanced Security Solution, Doha, Qatar

Abstract - Over the past decades, the malware complexity is constantly varying and increasing, which represents a severe security challenge. Artificial Intelligence (AI) based detection techniques are introduced for improving efficiency and accuracy of classification and detection rate of malwares in smart devices. One of the main limitations of AI-based detection is the low accuracy. It is critical to improve the accuracy and performance of malware detection systems to secure user data in smart devices from external threats. This study presents a comprehensive survey on modern AI techniques used for malware detection. The study provides an analysis for malware detection systems to identify the main components and different techniques used in each of these components. The study first identifies the main components in a detection system, then compares and analyses these components in recent research. The Data pre-processing, Feature extraction, and Malware detection (DFM) Taxonomy of three factors are proposed in this system to improve the detection classification rate for malware in smart devices. The proposed DFM taxonomy helps in the enhancement of smart devices' security. The taxonomy is evaluated on criteria of accurate detection and classification rate of malware detection. The performance is evaluated by comparing the proposed taxonomy with state of the art solution. This study provides deep analysis and discussion for the current state of art malware detection systems based on AI methods.

Adversarial of Defense Systems Against Spoofing Face Recognition Attacks

Ashokkumar Patel, Riley Kiefer
Department of Computer Science, Florida Polytechnic University, Lakeland, Florida, USA

Abstract— Face recognition and verification has been biometrics-based access and authentication system for several applications, including critical data access for aerospace software, next-generation air traffic control systems; aeronautical and space networks; airport and airline information systems; and commercial space vehicles. We have experimented, using a new innovative crafting approach, that by manipulating the deep feature representation extracted from the face image via imperceptibly small perturbations added at the pixel level, effectively bypass the defense methods for face recognition and verification system and consider fake one as original with a higher rate of success. One distinguishing feature of this innovated artificial image attack approach is that it is identity-agnostic; It behaves highly effective, even against identities unknown in advance. This specific approach shows that it can spoof almost all tested identities, including those whose identities are not known by forehand. So far results of our experiments indicate that a multiple-identity attack is a real threat and should be considered when deploying face recognition systems in practice.

Analysis of the Information Security of Public Organizations in Ecuador

Richard Romero Izurieta, Luis Jhony Caucha Morales, Segundo Moises Toapanta Toapanta,
Luis Enrique Mafla Gallegos, Jose Antonio Orizaga Trejo
National University of Tumbes, Peru;
State University of Milagro (UNEMI), Milagro, Ecuador;
Faculty of Engineering Sciences, Quevedo State Technical University (UTEQ), Quevedo, Ecuador;
Faculty of Engineering System, National Polytechnic School (EPN), Quito,Ecuador;
Department of Information System (CUCEA), University of Guadalajara, Guadalajara, Mexico;
School of Mathematical Computational, Sciences and Information Technology, Yachay Tech University, Ecuador

Abstract—Public organizations have the responsibility by law to manage sensitive data that must be shared; the problem was found that organizations are very vulnerable and that computer attacks are becoming more and more effective and dangerous. As an objective of this research, the Information Security of public organizations in Ecuador was analyzed, the current status was known and improvements in its management were proposed. The deductive method was applied for the review and analysis of factors and variables that allow improving information security management in public organizations. The result was an Information security management model based on strategic planning; process and matrices for the evaluation of the information security management capacity. It was concluded that public organizations in Ecuador have a low level of information security management capacity, 70% are at a "Formative" level and 22% at a "Managed" level. To Manage Information Security in an optimal way, it is necessary to start from a strategic planning that allows directing the resources and capacities available to the achievement of the objectives established for the organization.

Machine Learning for Multiple Stage Phishing URL Prediction

Khalid Amen, Mohamad Zohdy, Mohammed Mahmoud
Department of Electrical and Computer Engineering, Oakland University, Rochester, Michigan, USA

Abstract—Phishing is a fraudulent process and a form of cybercrime where an attacker tries to obtain sensitive information for malicious use. A phisher uses social engineering and technical deception to fetch private information from the web user. Previous Machine Learning (ML) approaches have been used to detect whether URLs are valid, or invalid. The purpose of this work is to detect, or predict, the three stages of Phishing URLs starting with valid, not enough info and invalid URLs. We will investigate different potential models that are trained by Machine Learning algorithms and find out which of these models has better accuracy.

Digital Twins for Industrial Control Systems Security

Guillermo Francia III, Gregory Hall
Center for Cybersecurity, University of West Florida, Pensacola, Florida, USA

Abstract—The emergence of Digital Twins in Industrial Control Systems has enabled advances in the test and evaluation of those systems in a low-cost and non-disruptive manner. The authors propose leveraging these benefits even farther, using the conformance of the digital twin to its physical counterpart to enable security testing of Industrial Control Systems for vulnerability assessment and penetration testing. A proof-of-concept system involving a bottle-filling system controlled by a Programmable Logic Controller (PLC) is presented.

Developing an API for Block-Cipher Encryption Powered by Supervised Learning

Atif Tirmizi, Osama Abuomar, Khaled M. Alzoubi
Department of Engineering, Computing, and Mathematical Sciences, Lewis University, Romeoville, Illinois, USA

Abstract—In cryptography, one of the standard classes of algorithms used today is called Advanced Encryption Standard (AES). Although these algorithms provide sufficient strength, they are not unbreakable, either by quantum computing or adversaries trying to decipher data. This study shows the potential of using machine learning (ML) by building an application Programming Interface (API) driven framework for encryption and decryption that utilizes ML to add abstracted layers of security over encrypted data. This abstracted data can only be interpreted by the developed framework. The utilization of this API will ultimately provide a cohesive system that can limit many vulnerabilities and prove how ML can be utilized as a tool against cyber-attacks.

Model-Based Security Testing of Vehicle Networks

Florian Sommer, Reiner Kriesten, Frank Kargl
Institute of Energy Efficient Mobility, Karlsruhe University of Applied Sciences, Karlsruhe, Germany;
Institute of Distributed Systems, Ulm University, Ulm, Germany

Abstract—Modern vehicles consist of a large number of electronic information technology components, which communicate with each other and external components. To protect vehicles against security attacks, automotive-specific standards and regulations require an integration of security concepts and measures in vehicles. Security testing techniques, such as penetration tests, are used to verify and validate those measures. However, these methods are usually carried out manually in late phases of development. Thus, identified vulnerabilities can only be eliminated at a late stage leading to a high investment of time and resources. This paper presents a model-based security testing approach which aims to enable security tests early on in the vehicle development process in an automated way. This allows vulnerabilities to be identified and eliminated at an early stage during development. Therefore, we show our concept to create a security model based on a vehicle network. This model can be used to automatically derive attack paths for security testing. We further illustrate our approach by applying it to a real-world vehicle network.

MalDeWe: New Malware Website Detector Model based on Natural Language Processing using Balanced Dataset

Jovana Dobreva, Aleksandra Popovska Mitrovikj, Vesna Dimitrova
Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje, R. N. Macedonia

Abstract—The increasing use of the Internet in everyday life and the huge number of users also leads to increasing the number of malicious websites, which aim is to damage a computer system or compromise data without the owner's consent. In this paper we propose a Natural Language Processing (NLP) model, called Malware Website Detector (MalDeWe), for malware web page identification that is trained on a domain-specific corpus. The major goal of this model is to transfer English word knowledge from the pre-trained model RoBERTa into a dataset of JavaScript codes that included the website's text context as well as certain JavaScript expressions. With this model we obtain a Roc Auc score of 0.95. Therefore, we can conclude that our model is doing admirably in terms of identifying a malicious web page. On the other hand, we may infer that one of the most essential aspects to consider while training a classification model is dataset balance. Whereas the model trained on the initial imbalanced dataset failed to detect harmful websites, the model, trained on the balanced dataset, correctly identified 95% of dangerous websites. In that sequence, we may deduce that the metrics used for the model evaluation are critical. Therefore, we recommend using the Roc Auc score, Recall, Precision, and Confusion matrix as evaluation metrics. So, in this paper we propose a new NLP model, and we discuss the reasons why the choice of evaluation metrics are important and how the dataset balance makes changes on the model efficacy.

Remote Video-Surveillance

Leonidas Deligiannidis
School of Computing and Data Science, Wentworth Institute of Technology, Boston, Massachusetts, USA

Abstract - We present a video surveillance solution that requires minimal installation and configuration effort. Video feeds are encrypted and accessible to anyone with the appropriate authentication credentials. The presented solution does not require any router nor firewall reconfiguration and the video feed is accessible even if the camera is behind a firewall and in a private network. Normally, port-forwarding needs to be enabled on the firewall, at minimum, to enable internet clients to access a camera located in a private network. Our solution enables a camera to be placed behind firewalls, in private networks, and still be accessible from the internet or other private networks without reconfiguring any of the existing network infrastructure. We achieve this by initiating a connection from the camera to a public proxy server which serves the role of an intermediary between clients/users and the camera. Clients connect to the proxy server. The proxy server then hands off the video feed of the camera to the client. This way, there is never a connection initiation from the internet to the camera. We provide three different methods of accessing the video feeds: via: a mobile app, a conventional browser, and a Linux command-line interface.

Developing Cybersecurity Workforce: Introducing Cybersec Labs for Industry Standard Cybersecurity Training

Yasin Kose, Murat Ozer, Mehmet Bastug, Said Varlioglu, Oguzhan Basibuyuk, Hari Priya Ponnakanti
Friedrich-Alexander-Universitat Erlangen-Nurnberg, Erlangen, Germany;
School of Information Technology, University of Cincinnati, Cincinnati, Ohio, USA;
Cybercrime and Homeland Security, University of Scranton, Scranton, PA, USA;
Homeland Security and Criminal Justice, Sul Ross State University, Alpine, Texas, USA

Abstract—In recent years, widespread internet usage, increasing numbers of IoT devices, and vast data on social media have increased cyber attack vectors. Parallel to this trend, the demand for the cyber security workforce has augmented; however, employers could not fill these positions due to the shortage in the field. This anomaly, in turn, has placed companies in a vulnerable situation by being exposed to an increased level of cyber security threats/incidents. Relevant authorities underscore the importance of a comprehensive cybersecurity workforce framework to overcome this problem. Given this context, this study offers a CyberSec Labs framework to improve the skills of potential cybersecurity actors. The framework extensively benefits from hands-on exercises like cloud computing and networking, OSINT, data-driven cybersecurity approach, vulnerability and penetration test analysis, and volatility analysis. The overall goal of this open-source material is to prepare students for their future cybersecurity roles, which contributes to overcoming the shortage of skilled human capital in the cybersecurity field.

First Considerations on How to Design a Honeypot to Protect (Swarms of) Drones by Fake Mass Deployment or by Using a Twin to Fake Interactions with the Physical World

Serge Chaumette, Titien Cubilier
Univ. Bordeaux, Bordeaux INP, CNRS, LaBRI, Talence, France;
Univ. Bordeaux, Talence, France

Abstract—Drones and swarms of drones are now considered an additional tool for both civilian and military applications. As any computer-based system they can thus be (and are) the target of attacks and the consequences of such attacks can be dramatic for assets and people. We believe an approach based on honeypots that would attract the attention of attackers and would behave so that these attackers could not even understand they are in a honeypot and not in a real drone, would be a significant step towards the protection of these systems. Even though some prototypes exist, they do not fully address the fact of luring the attacker to believe he/she controls a real drone. In this paper we present our work to address this issue.

Measuring the Effectiveness of Security Treatments on Threats through a Security ROI Model

Toby Houchens, Yanzhen Qu
School of Computer Science, Colorado Technical University, Colorado Springs, Colorado, USA

Abstract— With the fast expansion of the digital economy and more adoption of data driven decision making business management process in most industries, there is an urgent need for a quantitative model abstract enough to capture the complexity of security management, and practical enough to measure the effectiveness of any type of investment in security treatments towards business return on investment. In this paper we have developed a Security ROI (Return on Investment) model which is a function of three independent variables: (a) severity of security threat, (b) the probability of threat realization, and (c) the investment on security treatments. We have demonstrated the entire process of how the model is developed. Also, through a few examples that reflect some typical real-world investment on security treatments approaches, we have applied the Security ROI model to analyze the effectiveness of security treatments on threats in each of these scenarios to show the usability and suitability of the model.

Sprofiler: Automatic Generating System of Container-Native System Call Filtering Rules for Attack Surface Reduction

Takashi Iiguni, Hitoshi Kamei, Keizo Saisho
Graduate School of Engineering, Kagawa University, Kagawa, Japan;
Faculty of Engineering and Design, Kagawa University, Kagawa, Japan

Abstract—Containerized virtualization technologies are applied to many production environments. Compared with the VMs, the image size of containers is relatively small and the containers launch fast. Thus, containers can achieve fast scaleout, high portablity, and fast release cycles. Meanwhile, containerized virtualization sometimes causes security incidents when operated in a multi-tenant environment, such as public cloud. From previous researches, container runtimes may have a vulnerability of privilege escalation attack from a container to its host. One of the reasons is that the containerized virtualization shares the kernel of operating system among containers and their host. Even if there is no vulnerability in container runtimes, the privilege escalation attacks may be possible when the container runtimes are configured inadequately. As a countermeasure, the method of restricting system calls issued from containers can be applied. It is possible to restrict the system calls by using system call filter, such as Seccomp-BPF. However, it is difficult for system administrators to know the system calls issued by applications. Therefore, it is necessary to have a function that examines the system calls issued by the application and support the creation of filtering rules. In this study, we propose Sprofiler generates filtering rules that are suitable for a workload of a container by combining static analysis of application executables and dyamic analysis of system calls actually issued from the container. This paper describes the design and implementation of the Sprofiler. Moreover, it shows the evaluation results of the effectiveness by applying system call filtering rules generated by the Sprofiler.

The Need to Teach Security with Labs

Gary L. Wallace
Southern Utah University, Cedar City, Utah, USA;
(Retired) Commander and Director of Information Technology and Security, United States Navy, USA

Abstract - Recognizing the importance of Information Security in industry today, the relevant strategy literature offers several anecdotal and prescriptive narratives to strengthen certain educational concepts. The National Security Agency reports that employees are lacking sufficient knowledge in Information Security. The industry has an educational crisis of massive proportion and because of this; companies and countries are easily being electronically compromised at the cost of billions of dollars every year. Most of the current literature provides neither clear information nor an organizing framework to successfully tackle Information Security education in this region. This Capstone Report explains the outline and implementation taken to implement such curriculum at a higher educational institution.

A Study of the Landscape of Security Issues, Vulnerabilities, and Defense Mechanisms in Web Based Applications

Nick Rahimi
Department of Comuting Sciences & Computer Engineering, University of Southern Mississippi, Hattiesburg, Mississippi, USA

Abstract— web application is application software that runs on a web server, unlike computer-based software programs that are run locally on the operating system of the device. Web applications are accessed by the user through a web browser with an active network connection. Their existence has been enabled in the contemporary era of the multifaceted World Wide Web (WWW), whose advancement to incorporate varied components and technologies has affected the consistent development of web applications. Currently, 4.66 billion people are active Internet users as Worldometer stated. However, according to the latest data breach report by IBM, the cost of a data breach in 2021 is US$ 4.24 million, this is a 10% rise from the average cost in 2019 which was $3.86 million. This paper discusses the diverse security issues, vulnerabilities, and defense mechanisms associated with web applications. The paper illuminates a comprehensive landscape on evolving and emerging web application vulnerabilities and reveals ways to attack, evade, and detect pattern mechanisms within all crucial web threats.

Consideration of the Use of Smart Grid Cyberattacks as an Influence Attack and Appropriate Deterrence

Jeremy Straub
Institute for Cyber Security Education and Research, North Dakota State University, Fargo, North Dakota, USA

Abstract—Attacks on the 'smart' power grid are frequently discussed in the context of this being an attacker's goal. Attacks that occur inadvertently due to power grid equipment becoming infected by general purpose malware or targeted in a broader attack are also described. Attacks have also been discussed where the power grid is seen as a means to an end where the actual target is impacted or made vulnerable through the discontinuation of electrical service. Perhaps a more insidious form of attack is one where the power grid equipment is attacked and compromised, but the compromise is not actually used immediately. Instead, the threat of shutdown or damage is used as leverage to achieve another goal (and to discourage activities to try to remediate the vulnerability). This paper discusses these type of influence attacks. It considers what types of parties would utilize this type of attack and what purposes the attack would be used for. Several scenarios are presented and analyzed. From this analysis, the paper discusses the types of deterrence that would be effective and appropriate. General purpose pre-event deterrence strategies and response/retaliation strategies are both evaluated.

CSCI-ISNA:
Social Network Analysis, Social Media, & Mining

Determining the Motivations to Use Computer-Mediated Communication

Junghyun Nam
University of North Georgia, Georgia, USA

Abstract - The purpose of this research is to study the motivation attributes of Computer-mediated communication (CMC) technology use in general. This study seeks to address the issues by applying the uses and gratifications theory and attempts to explore the specific gratifications sought from the use of different CMC technologies. An empirical survey will be conducted, and factor analysis will be undertaken to extract the specific motivations to use each CMC technology. The results of this study may be applied to CMC technology design.

Optimized Ranking-based Community Detection

Matin Pirouz
California State University, Fresno, California, USA

Abstract - Detecting communities in complex networks gives rise to a new data mining challenge known as community ranking. Community ranking can be computed based on the influence of each community on the neighboring communities as well as on the entire network. Our main research question is to verify if there is a significant correlation between community ranking and popular nodes. We address the detection of significant and hidden communities by ranking them based on information flow in data networks. We define the influence of communities based on three key features. The first key feature is the how dense the relationships inside the community are while the second feature is how strong the intra-community links are. The third features extrapolates the centrality of a community in a given dataset. In this paper, several novel community ranker algorithms are introduced. These ranker algorithms quantify the detected communities for further ranking.

Uncertainty Estimation for Twitter Inference

Meng Hsiu Tsai, Nicole Marie Ely, Yingfeng Wang
Department of Computer Science and Engineering, University of Tennessee at Chattanooga, Tennessee, USA;
Department of Data Science and Business Analytics, Florida Polytechnic University, Lakeland, Florida, USA

Abstract—Twitter provides a platform for users to express their opinions in the form of Twitter messages, called tweets. Analyzing tweets from a specific city or demographical group requires geographical or demographical information. Unfortunately, most Twitter users do not provide these details in their profiles. To overcome this challenge, tools have been developed for inferring users' geographical and demographical information from Twitter data. Using inference results is risky due to the lack of uncertainty estimation of these results. Here, we present a framework to estimate uncertainties of Twitter inference results. The effectiveness of this framework is verified in experiments.

Who Discussed Further Global Warming in Social Networks? A Big Graph Analysis Perspective

Mauro Leon-Payano, Miguel Nunez-del-Prado, Hugo Alatrista-Salas
Pontificia Universidad Catolica del Peru, Lima, Peru;
Universidad de Ingenieria y Tecnologia, Lima, Peru

Abstract—The massification of social networks raises new types of opinion leaders, called influencers. An influencer can lead others to take actions, alter behaviors, or change their opinions. This phenomenon provides new opportunities for companies to communicate about their products or services through digital strategies. However, identifying influencers in social networks as Twitter is challenging due to the large amount and complex data. In the present work, a two-stage process is proposed for identifying opinion leaders. First, the PROV-DM model is used in order to build a large heterogeneous graph. Later, the Banzi algorithm quantifies the direct and indirect influence of each user using parallel processing. Finally, the PageRank algorithm is used as a baseline to compare our algorithm's performance using the influence diffusion metric.

Proposal of an Application for Sharing Daily Ambient Sound

Mizuki Watanabe, Takayuki Fujimoto
Information Sciences and Arts of Graduate School, Toyo University, Tokyo, Japan

Abstract—In this research, we propose a social media application for sharing daily ambient sound with others. Applications for sharing sounds such as Clubhouse or Spaces provided by Twitter have become quite popular in the world. Strictly speaking, these tools are not considered social media because they just give users a radio presenter experience. As this example shows, though there are many different types of social media today, some of them do not play the role of such media. In addition, ambient sound is generally considered something not to be shared with others. People are, however, inclined to share with others what they are not supposed to. In this research, we reiterate the original meaning of social media. We then propose a social media tool for sharing ambient sound in order to satisfy human wants.

An Adaptive Approach for Fake News Detection in Social Media: Single vs Cross Domain

Shubhangi Rastogi, Shabeg Singh Gill, Divya Bansal
Cyber Security Research Centre, Punjab Engineering College, Chandigarh, India;
Computer Science & Social Sciences, Indraprastha Institute of Information Technology, Delhi, India

Abstract—The extensive use of online information platforms over traditional news media has amplified the dissemination of fake news. Supervised machine learning based techniques are being extensively used in the detection of fake news in Social Media. However, the performance of such models degrades in the case of cross-domain data scenarios. In this study, we empirically show that the performance of a model depends on the domainspecific and agnostic case. To conduct this study, we extracted the tweets based on the Afghanistan crisis and developed a dataset which we call 'FakeBan'. The country has witnessed the sudden spread of misinformation where several actors are misusing it as ammunition, leading to far-flung troubling implications. We chose to study the most recent Afghanistan and experimented with three completely different domains widely involved in fake news: national crisis, healthcare, and politics. Several advanced datasets are already available in the domain of healthcare and politics. However, it takes a long time to build a labeled dataset based on a recent national crisis. We propose an adaptive fake news detection technique capable of selecting the model based on the domain (single or cross) and thus address the challenging issues arising from the voluminous and highly varied information available on social media. The results of our study affirm that in the case of domain-specific data, machine learning classifiers have performed well using a set of selected features out of twenty-one extracted features. In contrast, deep learning models, particularly the BERT model, have outperformed traditional machine learning classifiers in domain agnostic cases.

Time is Important in Fake News Detection: A Short Review

Shubhangi Rastogi, Divya Bansal
Cyber Security Research Centre, Punjab Engineering College (Deemed to be university), Chandigarh, India

Abstract—The prevalence of fake news has augmented with the rise of digital sources, especially social media. In this paper, current fake news research is studied and examined to offer a succinct road-map for future work. The survey presents a novel three-tier system depending on the lifespan of news and divides the research in three phases: early, mid and late-stage detection. The strategy to be followed for fake news detection varies with the time of detection. Fake news has shown adverse effects in a very short time period of propagation on social media. To mitigate this, it is required to detect fake news at an early stage when limited information about the news is available. In contrast, rich information can be examined like user engagement, propagation patterns, etc., at a later stage when news is deeply spread in the social network. Therefore, it is important to first analyze the time when the news disseminated, and then follow a suitable fake news detection methodology presented in the-state-of-the-art .

A Sentiment Analysis Approach for Abusive Content Detection using Improved Dataset

Aziz L. Ngou Njikam Abdou, Kun Zhang, Elie Fute Tagne
Nanjing University of Science and Technology, Nanjing, China;
Faculty of Engineeering and Technology, University of Buea, Buea, Cameroon

Abstract—The rapid growth of information and communications technologies has led to the generation of enormous amount of information daily. Consequently, there has been an increased interest in effective data processing. The nature of the information is varied and can in some cases include users' emotions and opinions. The specific case of abusive content is becoming more relevant these recent years leading to the proposition of many models which unfortunately suffer from the unbalanced nature of the dataset used. We propose a sentiment analysis approach which classifies social media posts according to three categories: hate, abusive and neutral. The approach is based on a constructed dataset which reduces unbalancing and improves classification results.

Living, Loving and Learning Online: Raising Awareness of the New Normal and its Security Considerations

Portia Buthelezi
Computer Science Department, University of South Africa, Florida, South Africa

Abstract— Social media use has infiltrated families and communities at large. Due to the continued rise in smart phone uptake and usage in developing countries, there is an ever rising social media acceptance and use. The situation with developing countries is that they sometimes fall behind in technology adoption, in relation to their developed counterparts. The adoption of social media comes with challenges of users offending each other and the need to regulate the social media platform environment. The regulation is implemented in the form of documented etiquette, and policies for acceptable usage. The rules of engagement. Social media users have been reported to have violated these rules of engagement and subsequently offend fellow social media users. Some of these offences have even dominated news channels in the form of public outcry about offensive statements posted on social media and what actions should be taken against the offender(s). This study investigates the reasons why the social media policies for acceptable use, continue to be violated and how this can be addressed. The study will follow an exploratory qualitative approach will be followed to identify the reasons for user non-compliance, and to determine the factors that affect user non-compliance and how to influence these factors towards compliance and non-offensive, collegial social media interactions. In the South African context, there has been an increase in hate speech cases expressed through social media. The findings of this study could be used by policy makers as guidelines for improved implementation of social media policies. The improved policy implementation could indicate to the social media users, if their social media post contains elements that could be offensive.

The Impact of Social Media on HCI

Daiyaan Ijaz, Fady Nissan, Henry Dare, Jonathan Williams, Sean Radatz, William Parker, Sherrene Bogle, Mohammed Mahmoud
Department of Computer Science and Engineering, Oakland University, Rochester, Michigan, USA;
Department of Computer Science, Humboldt State University, Arcata, California, USA

Abstract—The rising use of social media has changed how humans fundamentally interact with computers and with each other. The widespread and continuous use of social media has led to increasing awareness of its impact on mental health and the addictiveness of these interactions. It has also changed interpersonal interaction, how companies conduct marketing, and the way information is presented online. This paper discusses how the devices we use to communicate have become an integral part of life, how the increased connectivity has impacted our society, social media addiction, interface design, user experience and other developments with Human Computer Interaction (HCI) in social media.

Information Dissemination Model and Trend Forecast of Short Video Social Network

Jiawei Han, Limin Jiang, Ying Xu, Jingwen Yang, Taotao Sun
Department of Cyberspace Security, Changchun University, Changchun, China;
Department of Administration, Changchun University, Changchun, China

Abstract—With the advent of big data era, short video social network has gradually become an important battlefield for the outbreak of online public opinion. Therefore, the study for information transmission model is a hot topic for governing online public opinion. Based on improving the shortcomings of traditional SEIR model, we establish the BLSI model. Furthermore, the main affecting factors of information dissemination are selected by the principal component analysis method. Through changing the value of model parameters, the processes of information dissemination are simulated under different conditions. The simulation results are similar to the real network.

Effects of Different Recommendation Algorithms on Structure of Social Networks

Sepideh Banihashemi, Abdolreza Abhari
Department of Computer Science, Ryerson University, Toronto, ON, Canada

Abstract—In this paper, the effects of two algorithms (i.e., K-means clustering, Cosine similarity) on the structure of a social network (i.e., Twitter, scientific research papers) is studied to examine the formation of communities when the users follow the recommendations provided by a simulator. The relationship among the users can be either follower-followee in a Twitter dataset or paper-publication venue in a scientific research papers dataset. The purpose is to evaluate how much detected communities and the resulted network graph would be different when following the recommendation provided by the applied algorithms after system recommends top-N recommendations for a selected user.

Platform Neutrality and the Global Balance of Powers

Marcel S. Stolz
Department of Computer Science, Cybersecurity Analytics Group, Global Cybersecurity Capacity Centre, University of Oxford, Oxford, United Kingdom, UK

Abstract—This paper defines platform neutrality as a concept for large technology companies, most notably, social media platform providers. It is deduced from the concept of state neutrality, and acknowledges societal and political functions as well as state-like structures these companies have put into place. The paper argues that recent developments demonstrate a convergence of social media towards platform neutrality. It explains the benefit of platform neutrality both for businesses as well as societies.

Dynamic Model Change Analysis of Social Behavior Processes during the COVID19 Pandemic

Julian Pucheta, Sergio Laboret, H. Daniel Patino, Carlos Salas, Cristian Rodriguez Rivero, Martin Herrera
Facultad de Ciencias Exactas, Fisicas y Naturales, Universidad Nacional de Cordoba, Cordoba, Argentina;
Instituto de Automatica, Universidad Nacional de San Juan, San Juan, Argentina;
Facultad de Tecnologia y Ciencias Aplicadas, Universidad Nacional de Catamarca, Catamarca, Argentina;
Wiskunde en Informatica, Universiteit van Amsterdam, Amsterdam, The Netherlands

Abstract—An analysis of changes in dynamic process models described by variables that represent social behavior from the point of view of people's mobility and of economic indices in the framework of the Covid19 pandemic is presented. Here, the mobility described by Google and Apple is used as a proxy for the social behavior to correlate it with the dynamic evolution of daily Covid 19 infections. In addition, indices related from the global economy are used as a proxy of the socio-economic process, where two of ascending evolution (MSFT Microsoft and NASDAQ, Inc.) and another with smooth evolution (WTI oil gallon price) are analyzed. The evolution of such proxies are related to the daily Covid 19 cases. In the latter case, it is complex to detect a territorial region of influence given the number of origins of influences that the selected indices have, but the impact of the first peak in China and the subsequent evolution in the world can be studied, especially in our country and in the Netherlands. The main findings include that the underlying model for social behavior has changed in different stages, depending on the months of the year and that after mid-2021 an unstable equilibrium is on the track, with the addition of the new possibilities provided by the vaccination process and the rules of social coexistence. It is concluded that it is necessary to analyze what decision should be taken at the social level of public policy and what personal decisions for each individual.

CSCI-ISBD:
BIG DATA AND DATA SCIENCE

Mining the Impact of Social Networking on High-Frequency Financial Data

Ray Hashemi, Omid Ardakani, Chanchal Tamrakar, Jeffery Young
Department of Computer Science, Georgia Southern University, Savannah, Georgia, USA;
Department of Economoics, Georgia Southern University, Savannah, Georgia, USA;
Department of Marketing, Georgia Southern University, Savannah, Georgia, USA;
Department of Computer Sciecne, Clemson University, Clemson, South Carolina, USA

Abstract – TBA

Sentence Complexity

Ron Coleman
Computer Science Department, Marist College, Poughkeepsie, New York, USA

Abstract—In this paper study and test a novel hypothesis: namely, sentence complexity is a measure of readability. Our goal is not to replace nor necessarily improve readability tests but to investigate an experimental, data science application of word complexity. Word complexity is the number of words needed to define a word according to a reference dictionary. Sentence complexity is the aggregation or reduction of word complexity for words in a sentence. With the aid of natural language processing, we test our hypothesis using a corpus of 24 classic fairy tales. We chose these stories for their accessibility, popularity, and brevity. The data shows their reading difficulty to be centered around the 6^{th}-7^{th} grade level. The data also show the Flesch-Kincaid test and sentence complexity are negatively correlated in reading ease ($r=-0.31$, $P<0.01$) and positively correlated in grade level ($r=0.46$, $P<0.01$), respectively. That the correlations are moderate suggests sentence complexity is related to Flesch-Kincaid but not a proxy for it. Finally, the data further show that sentence complexity is weakly correlated with maximum number of syllables in a sentence ($r=0.12$, $P<0.01$), weakly correlated with maximum word length in a sentence ($r=0.20$, $P<0.01$), and moderately correlated with the number of words in a sentence ($r=0.52$, $P<0.01$). These findings support our hypothesis.

Visually Compatible Home Decor Recommendations Using Object Detection and Product Matching

Unaiza Ahsan, Yuanbo Wang, Alexander Guo, Kevin D. Tynes Jr., Tianlong Xu, Estelle Afshar, Xiquan Cui
The Home Depot, USA;
Georgia Institute of Technology, Atlanta, Georgia, USA

Abstract—Automatically recommending visually compatible products to Home Decor shoppers is a challenging task for ecommerce companies in the home improvement domain. However, few satisfactory solutions have been proposed to address this problem. In this paper, we propose a novel approach that uses a room scene image as the primary data source to generate visually compatible product recommendations. More specifically, we first detect products shown in the room scene image. Then, we use image retrieval techniques (e.g. color matching and triplet contrastive learning) to find the most similar products, if not the same, from the catalog. The system is designed to scale up to millions of products. To evaluate the performance of the proposed approach under various scenarios and use cases, we test it on several decor datasets, e.g. the catalog of a large home improvement retailer, a sizable public decor dataset, and customergenerated images posted on product reviews. We compare the approaches and determine that the triplet contrastive learning outperforms color matching for image retrieval. When tested live on the retailer's website, this new experience increases by 1.5% the user engagement and by 4.0% the Average Order Value (AOV) of transactions.

Atomic Layer Deposition Optimization Using Convolutional Neural Networks

Julian Cagnazzo, Osama "sam" Abuomar, Angel-Yanguas Gil, Jeffrey W. Elam
Department of Engineering, Computing, and Mathematical Sciences, Lewis University, Romeoville, Illinois, USA;
Applied Materials Division, Argonne National Laboratory, Argonne, Illinois, USA

Abstract—During this study, we trained a convolutional neural network to predict the optimal dosing time for novel atomic layer deposition (ALD) processes. Two types of models were generated. The single reaction model (SRM) makes predictions based on growth per cycle rates from a single ALD reaction. The multiple reaction model (MRM) makes predictions based on growth per cycle rates from ten examples of the same ALD process with different reagent dosing times. When evaluated on a validation set of novel ALD reaction profiles the SRM and MRM had RMSE of 0.9056 and 0.2678 respectively. The SRM and MRM had R2 -scores of 0.9044 and 0.9941 respectively.

Enhancing Textual Knowledge Discovery using a Context-Awareness Approach

Nour Matta, Nada Matta, Evan Giret, Nicolas Declerqc
LIST3N, University of Technology of Troyes, Troyes, France;
Data Department, Namkin, Troyes, France;
Namkin, Troyes, France

Abstract—Semantic relation extraction has been a known branch in Natural language processing allowing identification of relations between words. Knowledge-based approaches such as wordnet and pattern matching approaches such as Hearst for hypernyms are used along with others. The search for relations is limited to sentences which avert the rest of the text narrowing the semantic relation extraction. In this paper, we introduce the use of context for improving the extraction using global information from the text and domain ontology.

Predictive Analytics Approach based on Grey-Markov for Digital Economy

Guang Hu, Yue Tang
School of Computer Science, Fudan University, China;
School of Statistics and Information, Shanghai University of International Business and Economics,
Shanghai Key Laboratory of Data Science, Shanghai, China

Abstract—This paper presents a predictive analytics approach based on Grey-Markov for digital economy. It selects the cumulative value data of consumer goods retail sales in China from February to December 2020. Firstly, the GM(1,1) model is established to forecast retail sales of consumer goods for each month of 2020. Markov model is used to divide three transition states according to the prediction error. Then the cumulative values of February and March 2021 are predicted by GM(1,1), and the prediction results are corrected by Markov chain model. Through comparing the prediction results of GM(1,1) model and GM(1,1)-Markov model, it is found that the prediction accuracy of GM(1,1)-Markov model is higher, which indicates that the GM(1,1)-Markov model has certain practical significance for the short-term prediction of retail sales of consumer goods.

Network Intrusion Detection System Using Principal Component Analysis Algorithm and Decision Tree Classifier

Oyeyemi Osho, Sungbum Hong, Tor A. Kwembe
Department of Computational & Enabled Science and Engineering, Jackson State University, Mississippi, USA;
Department of Electrical & Computer Engineering and Computer Science, Jackson State University, Mississippi, USA;
Department of Mathematics & Statisyical Sciences, Jackson State University, Mississippi, USA

Abstract— Network Intrusion Detection Systems (IDS) have become expedient for network security and ensuring safety of all connected devices. Network Intrusion Detection System (IDS) alludes to observing network data information swiftly detecting any intrusion pattern and preventing any harmful effect of anomaly intrusion will cost the network. Traditional Intrusion Detection Systems have shown deficiency of effectively discovering new attacks because attack patterns need to be updated for the system to discover newer anomaly intrusion, this approach is a reactional style instead of actively monitoring the network for any attack before the harm is carried out. To combat this issue, a Network Intrusion Detection System (IDS) based on Principal Component Analysis (PCA) Algorithm was proposed, in this paper we present a supervised machine learning model to detect intrusion in the Network. Our strategy which is based on Principal Component Analysis (PCA) that works by protruding data elements onto a feature space which is an input data vector that encompass the notable variations among known data elements. The Principal Component Analysis (PCA) reduces the high dimensional data vectors and detection is administered in a low dimensional space with high efficiency and low usage of system resources. We rely on some experiments we perform over a Network Intrusion records from the CICIDS2017 dataset, first by direct application of these two algorithms on the raw data and secondly after projection of the dataset on the new feature space. The indexes of detection accuracy, detection time, precision rate, and recall rate were applied to evaluate the results.

Time Series Analysis of COVID-19 Cases in Humboldt County

Soeon Park, Mohammed Mahmoud, Sherrene Bogle
Humboldt State University, Arcata, California, USA;
Oakland University, Rochester, Michigan, USA

Abstract - The time series of COVID-19 daily cases in the U.S is analyzed by utilizing the county-level temporal data, from January 22, 2020 to October 18, 2021. Autocorrelation and partial autocorrelation show that time series of daily cases in Humboldt county has a 7-day seasonal pattern. Visualization and augmented Dickey-Fuller test show that time series of daily cases in Humboldt county is non-stationary. The seven-order difference reveals that the time series is stationary. There is a moderate positive correlation between daily cases and fully vaccination rate. Clustering analysis describes 33 counties have similar daily case pattern with Humboldt County by standard deviation of 0.003. This analysis can be used for future time-series forecasting and planning.

A Zero Trust Model Based Framework for Data Quality Assessment

Mahmood Mohammed, John R. Talburt
Information Science, University of Arkansas at Little Rock, Little Rock, Arkansas, USA

Abstract - Zero trust security model has been picking up adoption in various organizations due to its advantages. Data quality is still one of the fundamental challenges in data curation in many organizations where data consumers don't trust data due to associated quality issues. As a result, there is a lack of confidence in making business decisions based on data. We design a model based on the zero trust security model to demonstrate how the trust of data consumers can be established. We present a sample application to distinguish the traditional approach from the zero trust data quality framework.

Using Data Analytics to Forecast Violent Crime

Herious Cotton, Tor A. Kwembe
Computational Data Enabled Science and Engineering (CDS&E), Jackson State University, Jackson, Mississippi, USA;
Department of Mathematics and Statistical Sciences, Jackson State University, Jackson, Mississippi, USA

Abstract - In this paper, data analytics was used to analyze criminal data. Prophet model was used to predict homicide and rape in the Southeastern cities of Memphis Tennessee, Jackson Mississippi, and New Orleans Louisiana. In this paper the prophet model algorithm was used to show a correlation between violent crime and social economic conditions in Memphis Tennessee, Jackson Mississippi, and New Orleans Louisiana. However, The LSTM recurrent neural network model and the traditional neural network model have smaller RMSE. Thus, LSTM recurrent neural network model and traditional neural network model performed better than prophet and linear regression models. These promising outcomes will be significant to scholars, policymakers, and law enforcement officers.

A Modified Clustering Using Representatives (CURE) to Enhance and Optimize Tracking and Monitoring of Maritime Traffic in Real-time Using Automatic Identification System (AIS) Data

Cheronika Manyfield-Donald, Tor A. Kwembe, Jing-Ru C. Cheng
Computational and Data-Enabled Science and Engineering, Jackson State University, Mississippi, USA;
Department of Mathematics and Statistical Sciences, Jackson State University, Mississippi, USA;
Information Technology Lab., Engineer Research and Development Center,
U.S. Army Corps of Engineers, Halls Ferry, Vicksburg, Mississippi, USA

Abstract—In this paper, we introduce a modification of the Clustering Using Representatives (CURE) algorithm to enhance and optimize the tracking and monitoring of maritime traffic in real-time using the Automatic Identification System (AIS) data. In doing so, we present a 2-D points data collection system for the utilization of a modified unsupervised machine learning clustering method of the CURE algorithm integrated with a data streaming algorithm to develop a more efficient method that will directly assist in adverting accidents and support in vessels avoiding dangerous environments. Results are presented that show tracking in inland as well as open sea waterways.

Implementation of PCA Enabled Support Vector Machine using Cytokines to Differentiate Smokers versus Non-smokers

Seema Singh Saharan, Pankaj Nagar, Kate Townsend Creasy, Eveline O. Stock, James Feng, Mary J. Malloy, John P. Kane
Department of Statistics, University of Rajasthan, Jaipur, India;
University of California Berkeley Extension, California, USA;
University of California, San Francisco, California, USA

Abstract - Presently, the role of cytokines in severe illnesses associated with smoking, such as COPD, cancer, and cardiac disease, is being explored to enable mechanistic insights and potential intervention strategies. We are investigating the connection between levels of inflammatory plasma cytokines in smokers versus non-smokers. Disease indicator cytokines can be used to monitor the progression of disease which can help in the crucial task of prognosis and definitive diagnosis. Powerful and versatile Machine Learning algorithms can be leveraged to extract insights that cannot be obtained manually. We have applied Support Vector Machine on 65 plasma cytokines and other traditional biomarkers to differentiate smokers and nonsmokers. To optimize the classification separability, we have used the following techniques: Principal component analysis, 10-fold cross validation and variable importance. The primary metric of evaluation is AUC-ROC curve, though we have additionally recorded and compared prediction accuracy across classifiers. The results are very promising. The AUROC classification accuracy achieved by SVM using the selected predictor feature variables is .892 with a 95%CI(.853,.93). The most prominent cytokines / contributing to the classification, in the order of importance are: I-TAC, Age, TG, G-CSF-CSF-3, MDC-CCL22, Eotaxin-3, LIF, IL-2, Eotaxin-2, MIP-3alpha. Mechanism of disease and impact of indicators need/ to be explored more rigorously from the molecular aspect to direct future research and treatment pathways. AI and other ML algorithms, if used effectively, will accelerate the discovery and identification of disease indicators.

Prototyping: Sample Selection for Imbalanced Data

Edward Schwalb
Schwalb Consulting LLC, USA

Abstract—In the context of supervised learning, we are concerned with the task of identifying a subset of the example instances which maximize the predictive performance. In contrast to sampling, we do not generate new instances because we do not know how to reliably label them. We propose a simple and effective method with complexity that is linear in the size of the source data, and logarithmic in the size of the number of examples selected. We demonstrate empirically that very significant improvements are achievable on skewed data across a wide range of model types and data sets. In particular, we observe that the fraction achieving peak performance is proportional to the square root of the reciprocal of the skewness.

Similarity Based Methods for Faulty Pattern Detection in Predictive Maintenance

Somayeh Bakhtiari Ramezani, Logan Cummins, Brad Killen, Richard Carley, Shahram Rahimi, Maria Seale
Computer Science and Engineering, Mississippi State University, Starkville, Mississippi, USA;
Institute for System Engineering Research, Mississippi State University, Vicksburg, Mississippi, USA;
U.S. Army, Engineering Research and Development Center, Vicksburg, USA

Abstract—Detecting similarity between instances in large datasets is key to accurately predicting faulty patterns in Predictive Maintenance (PM). Most of the existing methods use the whole data to train a clustering or classification model, resulting in lower accuracy, especially in noisy data. Similaritybased methods (SBM) aim to increase the performance of datadriven algorithms by selecting the most similar training instances as the prototypes. This study evaluates the state-of-the-art SBMs used in the PM and summarizes the main challenges yet to be considered. Based on our observations, Dynamic Time Warping and Longest Common Sub Sequence are the most promising methods in PM.

Classification of Compton Camera Based Prompt Gamma Imaging for Proton Radiotherapy by Random Forests

Carlos A. Barajas, Gerson C. Kroiz, Matthias K. Gobbert, Jerimy C. Polf
Department of Mathematics and Statistics, University of Maryland, Baltimore County, Maryland, USA;
Department of Radiation Oncology, University of Maryland School of Medicine, Baltimore, Maryland, USA

Abstract—Proton beam radiotherapy is a method of cancer treatment that uses proton beams to irradiate cancerous tissue, while simultaneously sparing doses to healthy tissue. One promising method of real-time imaging during treatment is the use of a Compton camera, which can image prompt gamma rays that are emitted along the beam's path through the patient. However, because of limitations in the Compton camera's ability to detect prompt gammas, the reconstructed images are often noisy and unusable for verifying proton treatment delivery. We conduct a hyperparameter search to find an optimal random forest model. We then present the results of the best performing random forest model, which demonstrate that this ensemble method is less effective than competing machine learning techniques for this application.

Insurance Reserve Prediction: Opportunities and Challenges

Ayman Taha Farahat, Bernard Cosgrave, Wael Rashwan, Susan McKeever
Technological University Dublin, Dublin, Ireland; Cairo University, Egypt; DOCOsoft, Dublin, Ireland;
Technological University Dublin and CeADAR Technology Centre for Applied AI, Dublin, Ireland

Abstract—Predicting claims' reserve is a critical challenge for insurers and has dramatic consequences on their managerial, financial and underwriting decisions. The insurers' capital and their underwriting capacity of further business are subject to the unexpected reserve estimation. Increasing premium rates and adjusting the underwriting policy decisions may balance the impact of unexpected claims. Consequently, this will implicate their business opportunities negatively. Fortunately, several papers focusing on the prediction of insurance reserve have been published in the literature. In this paper, we provide a comprehensive review of the research on the insurance reserve prediction techniques in economics and actuarial science literature as well as machine learning and computer science literature. Moreover, we classify these techniques into different approaches based on the prediction mechanism they use in estimation. For each approach, we survey reserve prediction methods, and then show the similarities and differences among them. In addition, the review is armed with a discussion on main the challenges and the future opportunities.

Research on the Status of Industrial Sewage Equipment based on Clustering Ensemble

Qinlu Li, Tao Du, Rui Zhang, Jin Zhou, Shouning Qu
School of Information Science and Engineering, University of Jinan, Jinan, China;
Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan, China;
Shandong Management University, Jinan, China;
State Key Laboratory of High-end Server and Storage Technology, China

Abstract - The collected industrial data are usually unlabeled, and generally use clustering algorithm for data analysis. However, a single clustering algorithm can not guarantee a good analysis of industrial datasets, and there are many limitations for users during processing. By fusing the multiple results obtained by the analysis, clustering ensemble can get the results of datasets more effectively. Therefore, we use density-based selective clustering ensemble method to analysis the collected datasets. The main steps are as follows: (1) three density based clustering algorithms are used to analyze the datasets; (2) in order to remove the negative influence of noise outliers, the elements in the obtained co-association matrix are zeroed for many times, and the corresponding results are obtained by hierarchical clustering; (3) set the internal validity index to select the results to get the best result. Through the experiment and the feedback of relevant factories, the effectiveness of the experiment is proved.

Data Visualization Tool for Covid-19 and Crime Data

Sean Walker, Sharad Sharma
Department of Computer Science, Bowie State University, Bowie, Maryland, USA

Abstract— Data visualization gives a visual context through maps or graphs and makes it easier for the human mind to identify trends, patterns, and outliers within large data sets. The understanding of patterns and location of crime through data visualization and data mining techniques approaches is a very useful tool which can help and support police forces. Identification of crime characteristics and types are the first step for developing further analysis. This paper describes the development of data visualization tool using Unity 3D and Maptitude GIS for visualization of Baltimore COVID-19 and crime data. This effort aims to determine parameters that influence the vulnerability of African Americans to COVID-19 during the pandemic. The study has found that the factors shown to be influential in a person's susceptibility include neighborhood and physical environment, housing, occupation, education, income, and wealth gaps. The data collected from Baltimore incident reports and findings shared by Maryland SOA office shows that crime has increased and decreased in different areas during the time of COVID pandemic.

Data Analysis of Crime and Rates of Hospitalization due to COVID-19

Tyren Walker, Sharad Sharma
Department of Computer Science, Bowie State University, Bowie, Maryland, USA

Abstract—There has been an increasing concern that African American community has been disproportionately impacted during the coronavirus pandemic. This paper analyzes why the African American community is disproportionately impacted during the coronavirus pandemic and compares the COVID-19 data with hospitalizations, real estate, school closings, and crime data. Human behavior was impacted as a result of lockdown due to COVID pandemic and it lead to a shift in crime dynamics. We analyze shifts in crime types by comparing crimes before and after the COVID pandemic in Baltimore. There was a significant decline in total crimes during the time period immediately following stay at home orders. Findings show that the disproportionality among the African American community is significantly influenced by factors such as living in more crowded housing situations, working in consumer-facing serviced industries, having higher rates of preexisting medical conditions, and lack of insurance or a consistent care source.

CSCI-ISAI: ARTIFICIAL INTELLIGENCE

LogicGAN-based Data Augmentation Approach to Improve Adversarial Attack DNN Classiffiers

Christophe Feltus
Luxembourg Institute of Science and Technology (LIST), Luxembourg

Abstract—This position paper presents an innovative algorithmic approach in order to improve adversarial attack classifiers, based on data augmented by minor modifications generated by a logicGAN. Therefore, the paper addresses a particular type of mitigation against adversarial attacks, which consists of training the ”attacked” classifier with initial and adversarial data already known by the defender. Accordingly, we propose an algorithm that improves the training of the classifier: (1) by generating complementary adversarial data which instead of coming from the known adversarial attack, comes directly from minor modifications resulting from the already known adversarial data, and (2) by generating these minor modifications using a specific kind of generative adversarial network named logicGAN. By using an xAI system, this derivative of GAN has the particularity of yielding more substantial corrective feedback from the discriminator to the generator and, thereby, making the mitigation of adversarial attacks faster.

The Effects of Different Parameters on the Accuracy of Deep Learning Models for Predicting U.S. Citizen's Life Expectancy

Michelle Hu, Yen-Hung (Frank) Hu
Woodside High School, Newport News, Virginia, USA;
Department of Computer Science, Norfolk State University, Norfolk, Virginia, USA

Abstract—The increasing prevalence of deep learning-based machines in the daily life of average people results in a demand for research to be done on improving the accuracy of deep learning models. In response to this need, this paper aims to explore the effects of changing the parameters of a deep learning model, including the neuron count, epoch count, batch size, and validation split on the prediction accuracy of a deep learning model. We used the programming language Python, the TensorFlow and Pandas libraries, and the Keras application programming interface to create 13 regression-based deep learning models, all but one, which was used as a standard, of which had a parameter altered to be lower or higher than the standard model. After training each model using a dataset comprised of data from the 2010 United States census, we measured the predictive accuracy of each model at different epoch counts using the absolute average difference between the predictions of life expectancy from the models and the actual value from the 2010 U.S. census dataset. Based on the absolute average difference for each model, we found that increasing the neuron count, epoch count, and batch size and decreasing the validation split improves prediction accuracy in deep learning models, in most cases. These results can be used to create more accurate deep learning models for scientific or commercial use, and the models themselves can be used for their ability to predict life expectancies from given data, based on learned trends.

ConvGRU-TSNet: A Novel Deep Learning Approach for Multivariate Time Series Forecasting

Jingke She, Shanshan Gong, Suyuan Yang, Tianzi Shi, Shaofei Lu, Heng Wu
College of Computer Science and Electronic Engineering, Hunan University, Changsha, China;
Department of Mathematics and Computer Science, West Virginia State University, Morgantown, West Virginia, USA

Abstract—A novel deep learning approach, using Convolutional Gated Recurrent Unit with Temporal Spatial Neural Network (ConvGRU-TSNet), is presented in this paper. It is proposed to solve the mixed pattern challenges faced by multivariate time series forecasting. GRU and Convolutional GRU (ConvGRU) are utilized to extract long-term patterns among different periods such that the short-term patterns in one period can be detected. Additionally, the traditional autoregressive model is employed to tackle the scale insensitive problem. The experiments are carried out using four datasets with complex mixed patterns. It is observed that the ConvGRU-TSNet represents enhanced performance exceeding several baseline methods. Among the four selected datasets, the Relative Squared Error (RSE) can be reduced by 7.604% at most and the Correlation Coefficient (CORR) can be increased by 1.198% at most. It is demonstrated that the proposed ConvGRU-TSNet offers an altenative perspective of solving the mixed-pattern challenges with improved forecasting accuracy.

A Deep Learning BERT-Based Approach to Person-Job Fit in Talent Recruitment

Elias Abdollahnejad, Marilynn Kalman, Behrouz Far
Department of Electrical & Software Engineering, University of Calgary, Calgary, Canada;
Hireground Software Solutions, Calgary, Canada

Abstract—Although the widespread use of the Internet provides job recruiters with a larger pool to select the most qualified candidates, the tedious process of going over hundreds of resumes makes a fair and objective decision making more difficult. This paper proposes an end-to-end BERT-based framework to decrease the workload and expedite the shortlisting process of job applicants. Utilizing historicalrecords data of thousands failed and successful job applications, our model simulates the recruiters' decision-making process by the state-of-the-art BERT algorithm. The results show that BERT outperforms a variety of models by a high margin.

Attitudes Toward Corona Vaccine and Intention Among University Students

Sanaa Abujilban, Lina Mrayan, Shaher Hamaideh, Dheaya Alrousan, Hashim Abu-Gellban
Hashemite University, Jordan;
Texas Tech University, Lubbock, Texas, USA

Abstract— In this paper, we used a cross-sectional (correlational design) and snowballing sampling to analyze students' attitudes toward vaccination to identify its relation to their intentions to take the vaccine. We applied a multiple linear regression algorithm to predict the relationship between attitudes and intentions to take the vaccine. We found that participants were having intermediate to high mistrust of vaccine benefits and high worries over unforeseen future effects. Furthermore, we found that only two independent variables (mistrusts of vaccine benefits and the number of flu vaccination) contributed significantly to the prediction of the intention to take the vaccination level.

Pairwise Learning for Imbalanced Data Classification

Shu Liu, Qiang Wu
Computational and Data Science PhD Program, Middle Tennessee State University, Murfreesboro, Tennessee, USA;
Department of Mathematical Sciences, Middle Tennessee State University, Murfreesboro, Tennessee, USA

Abstract—Imbalanced data classification problems appear quite commonly in real-world applications and impose great challenges to traditional classification approaches which work well only on balanced data but usually perform poorly on the minority class when the data is imbalanced. Resampling preprocessing by oversampling the minority class or downsampling the majority class helps improve the performance but may suffer from overfitting or loss of information. In this paper we propose a novel method called pairwise robust support vector machine (PRSVM) to overcome the difficulty of imbalanced data classification. It adapts the non-convex robust support vector classification loss to the pairwise learning setting. In the training process, samples from the minority class and the majority class always appear as pairs. This automatically balances the impact of two classes. Simulations and real-world applications show that PRSVM is highly effective.

Artificial Intelligence-Based Cognitive Radar Architecture

Arkadiusz Czuba
Warsaw University of Technology, Warsaw, Poland;
PIT-RADWAR S.A., Warsaw, Poland

Abstract—Cognitive architectures are part of artificial intelligence. They have found an application in among other places, robotics and autonomous driving. However, research related to the topic of integrating artificial intelligence architectures with cognitive radar is limited. In this paper, cognitive abilities such as learning, perception, attention, and decision-making mechanisms were conformed to radar capabilities. As a result of that customization, a novel cognitive radar architecture has been proposed. This concept is introducing promising multimodal perception for MultipleInput-Multiple-Output (MIMO) techniques, visual attention mechanisms for resource allocation and target prioritization, memory-based cognitive models, and reinforcement learning for learning receiver-transmitter relation. All of the aforementioned components combined look promising to improve overall radar performance.

Development of Algorithm to Predict Arrhythmias based on Ensemble Network using Electrocardiogram

Siho Shin, Mingu Kang, Gengjia Zhang, Jaehyo Jung, Meina Li, Youn Tae Kim
Department of IT Fusion Technology, AI Healthcare Research Center, Chosun University, Dong-gu, Gwangju, Korea;
College of Instrumentation and Electrical Engineering, Jilin University, Jilin, China

Abstract—Professional drivers are known to suffer from negative effects on their physical and mental health, which may even result in death due to the professional hazard of long working hours. For example, if a worker suffers from arrhythmia and loses consciousness while working, it can result in a secondary major accident, which is a social problem. In addition, since arrhythmias directly impact life, early detection and treatment are crucial. Therefore, we propose a CNN-LSTM-based ensemble network that can diagnose arrhythmias for professionals that work long hours, in real time during daily life. The proposed model includes a filter that removes noise for each frequency band in the system using wavelet transform and achieves 98% arrhythmia detection accuracy through an AI-based ensemble algorithm. In the future, our system can collect multiple biosignals to detect extensive disease information, thus achieving real-time health management and social cost reduction for professionals that work long hours.

Enhancing the Resolution of Satellite Imagery using a Generative Model

Mahee Tayba, Pablo Rivas
Department of Computer Science, Baylor University, Texas, USA

Abstract—Recent breakthroughs in deep learning algorithms introduced the image super-resolution technique that maps the low-resolution image to generate a high-resolution image. These techniques increase various surveillance applications by providing finer spatial details than data from original sensors. Satellite images obtained from Moderate Resolution Imaging Spectroradiometer (MODIS) observation offer essential information about the earth's landscape, ocean, and ecosystem, contributing to monitoring various applications in the scientific field. The spatial resolution of satellite images has a significant impact on image accuracy. This paper focuses on improving image resolution by training a convolutional neural network to produce super-resolution images from low-resolution images. We present an implementation of Super Resolution Generative Adversarial Network (SRGAN), a GAN-based approach that uses a perceptual loss function that includes an adversarial loss and a content loss. Using a discriminator network that is designed for discerning between super-resolved images and original photorealistic images, the adversarial loss drives the solution of this architecture to natural images. Moreover, the content loss is driven by perceptual similarity rather than pixel space similarity. We used this architecture to satellite images collected from NASA MODIS devices and found satisfactory results. Our key finding is that our system's result can now be used to improve a variety of low-resolution images.

Human Activity Classification Using Basic Machine Learning Models

Bikram Khanal, Pablo Rivas, Javier Orduz
Department of Computer Science, Baylor University, Texas, USA

Abstract—Human activity recognition (HAR) is the object of interest for many researchers in machine learning. In principle, providing accurate and reasonable information on an individual's activities and movements for pervasive computing is a very challenging problem. Recent advances in HAR have led to advanced tracking of highly complex human behaviors. This is progressively driving humans and computers to become seamlessly integrated through devices and software. The impact of this type of research has numerous applications in different sectors. This paper presents our initial experiments on evaluating the performance of popular machine learning algorithms in predicting human behaviors accurately. Our experiments suggest that some models can accomplish high recognition accuracy and low computational cost.

Quantifying the Right Combination of Knowledge and Population Diversity for Transferred Populations

Tami Alghamdi, Robert B. Heckendorn, Terence Soule
Department of Computer Science, University of Idaho, Moscow, Idaho, USA

Abstract—Transfer Learning (TL) is a process of leveraging knowledge from one problem called the source problem (S) to a related problem called the target problem (T). Experiments and previous research (see Section II) showed that the knowledge of solving a related problem and population diversity are able to aid machine learning algorithm to solve a related problem more easily. To understand the right amount of knowledge and diversity of the transferred population, we modeled the TL process using Evolutionary Computation (EC). Our model gives us more access and control over the TL components, especially the content of the transferred population. The results showed that the relationship R between the source S and the target T problems can quantify the right combination of each component of the transferred population.

An Altmetric Study of Artificial Intelligence in Medicine

David Wu, Qiping Zhang
Great Neck North High School, Great Neck, New York, USA;
Palmer School of Library and Information Science, Long Island University, Brookville, New York, USA

Abstract— Artificial Intelligence (AI) and its powered technologies have been crucial for medical research and clinical practice. While AI in medicine is a rapidly evolving field, this study provides a new perspective to review its related bodies of literature through a bibliometric analysis of Altmetric data. Altimetric is a system that tracks the attention that research outputs such as scholarly articles receive from different web sources (e.g. social media, mainstream news, blogs). Altmetric Explorer is an online platform that enables users to browse and report on citation-based attention metric data for a given scholarly output (including journal articles and dataset). The purpose of this study is to perform an Altmetric analysis to systematically study research trends on AI in medicine. The study identifies various aspects of research outputs on AI in medicine (mentions, attention scores, a timeline of mentions, a timeline of yearly publication, top 9 journals, top 13 affiliatio1n of first authors, and twitter demographics). These findings would be of interest to both researchers and practitioners in the field. In conclusion, research of AI in medicine has been attracting significant attention in the past 4-5 years and top journals and top research affiliations follow the trend of Bradford's law and Lotka's law.

A Study on Deep Learning Approach to Optimize Solving Construction Problems

Phillip Roshon, Feng-Jen Yang
Department of Computer Science, Florida Polytechnic University, Lakeland, Florida, USA

Abstract—In this study, we focus on a problem domain, construction problems, for reinforcement learning systems to optimize. We relate our approach to existing research in the field of automated theorem proving and other related techniques to optimize the solutions in this domain. We expect this study can inspire more interest in the adoption of and improve the efficiency of existing production systems.

Stable Parallel Training of Wasserstein Conditional Generative Adversarial Neural Networks

Massimiliano Lupo Pasini, Junqi Yin
Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA;
National Center for Computational Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Abstract—We use a stable parallel approach to train Wasserstein Conditional Generative Adversarial Neural Networks (WCGANs). The parallel training reduces the risk of mode collapse and enhances scalability by using multiple generators that are concurrently trained, each one of them focusing on a single data label. The use of the Wasserstein metric reduces the risk of cycling by stabilizing the training of each generator. We apply the approach on the CIFAR10 and the CIFAR100 datasets, two standard benchmark datasets with images of the same resolution, but different number of classes. Performance is assessed using the inception score, the Frechet inception distance, and image ´ quality. An improvement in inception score and Frechet inception ´ distance is shown in comparison to previous results obtained by performing the parallel approach on deep convolutional conditional generative adversarial neural networks (DC-CGANs). Weak scaling is attained on both datasets using up to 100 NVIDIA V100 GPUs on the OLCF supercomputer Summit.

An Online Reinforcement Learning Approach for Solving the Dynamic Flexible Job-Shop Scheduling Problem for Multiple Products and Constraints

Nour El-Din Ali Said, Yassin Samaha, Eman Azab, Lamia A. Shihata, Maggie Mashaly
German University in Cairo, Egypt

Abstract - In the manufacturing industries, the most challenging problems are mostly related to time efficiency and customer satisfaction. This is mainly translated to how efficient is the frequent task of scheduling jobs to alternative routes on a number of machines. Although scheduling has been studied for decades, there is a shortage to a generalized approach for the production scheduling that can adapt to changes occurring continuously at any production environment. This research work addresses the dynamic production scheduling problem and the optimization techniques that could be applied to the production schedule to increase its efficiency. An algorithm is proposed to apply the Q-learning optimization technique on a dynamic flexible job-shop scheduling problem of a real case study of a pharmaceutical factory with 18 machines and 22 products. Proposed algorithm is shown to be able to achieve an efficient schedule with short make-span in minimal time duration and without requiring any learning process from previous schedules, thus increasing the factory's overall efficiency. In addition, the proposed algorithm operates online as any change occurring in the production environment is signaled automatically to it where it responds be regenerating the most optimal updated production schedule.

PCA Approaches for Optimal Convolution Kernels in Convolutional Neural Networks

Philku Lee, Deyeon Kim, Seung Heon Lee, Seongjai Kim
Center for Advanced Vehicular Systems, Mississippi State University, Mississippi State, MS, USA;
Chase Bank, USA;
Department of Biochemistry and Molecular Biology, Mississippi State University, Mississippi State, MS, USA;
Department of Mathematics and Statistics, Mississippi State University, Mississippi State, MS, USA

Abstract—Convolutional neural networks (CNNs) have become one of most powerful machine learning models; with enough data, their accuracy in tasks such as image-related classifications and natural language processing is unmatched. The drawback that many scientists have commented on is the fact that these networks, usually trained from randomly-initialized parameters, are black-boxes. This article introduces an innovative variant for CNNs, which incorporates principal components (PCs) derived from well-trained convolution kernels. The variant is called the principal component-incorporating CNN (PC-CNN), in which the PCs are employed either as a complete replacement for randomlyinitialized convolution kernels or as an initialization for the convolution kernels to be re-trained. The objective is to help training processes converge to the global minimizer. The PC-CNN is applied for the MNIST handwritten digit dataset to prove its effectiveness.

A Federated Affective Computing Framework to Learn from Small Data

Luca Bondin, Alexiei Dingli
Department of Artificial Intelligence, University of Malta, Malta

Abstract— The field of Affective Computing has for the past decade transformed into one of the most vibrant areas of research under the domain of Artificial Intelligence. However, the development of large-scale affect-enabled applications has been held back to the data required to train models for affect recognition. This data often does not exist or has its access restricted due to privacy-related issues if it does exist. We propose introducing concepts such as Federated Learning to the area of Affective Computing to overcome these shortcomings. Federated Learning will allow it to source data from many separate agents, helping train models faster while preserving their privacy. We introduce the Federated Affective Computing framework, a Composite AI framework, that improves traditional affective computing approaches' performance. More importantly, it aims to make large-scale affect-enabled applications feasible, even in scenarios where limited data is available to create such solutions. From evaluation that has been carried out on a Federated Affective Computing case study, we prove that the creation of intelligent models is made possible even when large enough bodies of data to train such models do not exist. We also show that we make possible the process of "continuous learning" while reducing the risk of catastrophic for-getting through our implementation.

On the Practical Uses of Experimental Adversarial Neural Cryptography

Korn Sooksatra, Pablo Rivas
School of Engineering and Computer Science, Department of Computer Science, Baylor University, Texas, USA

Abstract—With the rise of generative adversarial networks (GANs), many areas have seen remarkable improvements, e.g., computer vision, natural language processing, and the medical field. Notably, cryptography has been fueled by GANs producing what is called adversarial neural cryptography (ANC). However, in these five years, ANC has little documented experimentation and applications that can be used in the real world. This paper aims to perform experiments on ANC to verify if the current status of ANC is ready for practical implementations of symmetric-key encryption. In our investigation, we assess several entities in ANC during training, encryption, and decryption time of an ANC model, including decryption accuracy analysis. Furthermore, we study the resources required for deployment using different quantization techniques to reduce the size of an ANC model and its impact on performance and encryption accuracy. Our study provides enough data for offering practical advice for using and implementing ANC models.

Weakly Supervised Crowdsourcing Learning Based on Adversarial Consensus

Meng-Long Wei, Shao-Yuan Li, Sheng-Jun Huang
Ministry of Industry and Information Technology Key Laboratory of Pattern Analysis and Machine, Intelligence College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China

Abstract—Crowdsourcing provides an efficient way to obtain labels for large datasets in the deep learning era. However, due to the non-expert workers, the annotations are usually noisy. Besides, concerning the labeling cost, sparse annotations are common. To face this challenge, we propose an approach based on adversarial consensus, which trains one classifier for each worker, and enforces their predictions over the ground-truth labels to be maximally consistent by exploiting the generative adversarial learning idea. We give two implementations respectively for the light and heavy noise cases. Extensive experiments on real-world and synthetic datasets demonstrate the effectiveness of our approach.

Adversarial Training Negatively Affects Fairness

Korn Sooksatra, Pablo Rivas
School of Engineering and Computer Science, Department of Computer Science, Baylor University, Texas, USA

Abstract—With the increasing presence of deep learning models, many applications have had significant improvements; however, they face a new vulnerability known as adversarial examples. Adversarial examples can mislead deep learning models to predict the wrong classes without being noticed by human actors. Recently, many works have tried to improve adversarial examples to make them stronger and more effective. However, although some researchers have invented mechanisms to defend deep learning models against adversarial examples, those mechanisms may negatively affect different measures of fairness, which are critical in practice. In this work, we mathematically define four fairness scores to show that training adversarially robust models can have a negative effect on fairness scores. Furthermore, we empirically show that adversarial training, one of the most potent defensive mechanisms against adversarial examples, can harm them.

Application of Data Transformation Techniques and Train-Test Split

Amaya Shepard, Naima Naheed
Computer Science, Physics and Engineering, Benedict College, Columbia, South Carolina, USA

Abstract—With the increasing presence of deep learning models, many applications have had significant improvements; however, they face a new vulnerability known as adversarial examples. Adversarial examples can mislead deep learning models to predict the wrong classes without being noticed by human actors. Recently, many works have tried to improve adversarial examples to make them stronger and more effective. However, although some researchers have invented mechanisms to defend deep learning models against adversarial examples, those mechanisms may negatively affect different measures of fairness, which are critical in practice. In this work, we mathematically define four fairness scores to show that training adversarially robust models can have a negative effect on fairness scores. Furthermore, we empirically show that adversarial training, one of the most potent defensive mechanisms against adversarial examples, can harm them.

Modeling SQL Statement Correctness with Attention-Based Convolutional Neural Networks

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

Abstract—Automated grading of SQL statements is a topic of interest for instructors and students alike. It can give teachers additional time to be more effective in identifying issues quickly, and it can give students a preview of the grade they may receive. Existing attempts in the literature create models based on a variety of methodologies that exploit the structure of SQL statements and model answers; however, very few have leveraged the recent advances in deep learning. This paper employs a convolutional self-attention mechanism to learn complex contextual and grammatical dependencies directly from data of labeled SQL statements. Our experiments suggest that the proposed parameter-sharing strategy can adequately model the problem of detecting the correctness of an SQL statement with a balanced accuracy of up to 81.2% and an AUC of 0.87 in cross-validation.

Face of the Team - Diversity, Equity, and Inclusion

Tsz Hin (Martin) Cheung, Erik Noyes, Leonidas Deligiannidis
College of Computing and Data Science, Wentworth Institute of Technology, Boston, Massachusetts, USA;
Blank Center for Entrepreneurship, Babson College, Babson Park, Massachusetts, USA

Abstract - Among the many breakthroughs in machine learning, and specifically generative deep learning, trained neural networks can render entirely synthetic, photorealistic faces -- reflecting a deep understanding of face compositions and variability. This deep understanding and rendering power can be harnessed for powerful social commentary about Diversity, Equity, and Inclusion (DEI) in the business landscape, and support A.I.-enabled activism about inequalities in corporate leadership, in venture finance, and the overall success and failure of DEI initiatives. Properly trained generative deep learning models can enable the interpolation of latent space/latent dimensions in facial structures, thus enabling the hybridization of management team facial attributes to depict team diversity in an "averaged" team face. This project leverages state of the art, pre-trained generative face-creation models to support multi-variable interpolation between submitted faces, e.g., gender, race, age. The contribution is to show a powerful new way to track the (non)diversity of powerful management teams.

Quantum Machine Learning: A Case Study of Grover's Algorithm

Bikram Khanal, Pablo Rivas, Javier Orduz
School of Engineering and Computer Science, Department of Computer Science, Baylor University, Texas, USA

Abstract—The complexity of searching algorithms in classical computing is a perpetual researched field. Quantum computers and quantum algorithms can compute these problems faster, and, in addition, machine learning implementation could provide a prominent way to boost quantum technology. We call quantum machine learning to this novel set of tools coming from artificial intelligence and quantum mechanics. To achieve our purpose, we focus on applications on quantum machine learning; in particular, we propose a review and exploration of topics such as variational quantum algorithms, kernel methods, and a review of Grover's algorithm (GA) as a quantum classifier. We start with the GA exploration to achieve this goal, which is a quantum search algorithm that achieves quadratic speedup over optimal classical search implementation. This paper implements a GA exploration that includes a concept summary and implementation, considering only AND, XOR, and OR gates. We also discuss potential in quantum machine learning.

Evaluating Accuracy and Adversarial Robustness of Quanvolutional Neural Networks

Korn Sooksatra, Pablo Rivas, Javier Orduz
School of Engineering and Computer Science, Department of Computer Science, Baylor University, Texas, USA

Abstract—Machine learning can thrust technological advances and benefit different application areas. Further, with the rise of quantum computing, machine learning algorithms have begun to be implemented in a quantum environment; this is now referred to as quantum machine learning. There are several attempts to implement deep learning in quantum computers. Nevertheless, they were not entirely successful. Then, a convolutional neural network (CNN) combined with an additional quanvolutional layer was discovered and called a quanvolutional neural network (QNN). A QNN has shown a higher performance over a classical CNN. As a result, QNNs could achieve better accuracy and loss values than the classical ones and show their robustness against adversarial examples generated from their classical versions. This work aims to evaluate the accuracy, loss values, and adversarial robustness of QNNs compared to CNNs.

Hybrid Quantum Variational Autoencoders for Representation Learning

Pablo Rivas, Liang Zhao, Zhengrong Zhang, Javier Orduz
School of Engineering and Computer Science, Department of Computer Science, Baylor University, Texas, USA;
Marketing Department, St. Ambrose University, Davenport, Iowa, USA;
College of Arts & Sciences, Department of Physics, Baylor University, Texas, USA

Abstract—Representation learning is a standard area that has seen many improvements based on machine learning advances. Quantum machine learning advances are now spreading across different application areas such as representation learning. This paper introduces a novel hybrid quantum machine learning approach to representation learning by using a quantum variational circuit that is trainable with traditional gradient descent techiniques. We use marketing data to showcase the learning potential of our model.

PM2.5 Prediction Based on a Combined EMD-LSTM Fusion Model

Jingyi Zhao, Fahu He, Zhanlin Ji, Ivan Ganchev
College of Artificial Intelligence, North China University of Science and Technology, Caofeidian, Tangshan City, China;
Telecommunications Research Centre, University of Limerick, Limerick, Ireland;
University of Plovdiv "Paisii Hilendarski", Plovdiv, Bulgaria;
Institute of Mathematics and Informatics - Bulgarian Academy of Sciences, Sofia, Bulgaria

Abstract—Most of the research, conducted to date on the prediction of Fine Particulate Matter with a diameter less than 2.5 micrometers (PM2.5), based on machine learning and deep learning techniques, ignores the fact that the PM2.5 values are constantly changing over time. Although many researchers use Long Short-Term Memory (LSTM) neural networks based on time series to predict PM2.5 values, due to the instability of data, the results often had a certain lag. This paper proposes a combined Empirical Mode Decomposition (EMD)–LSTM fusion model for the prediction of PM2.5 values. To evaluate the performance of the model in comparison to other existing models, experiments were conducted with a public PM2.5 data set, using the root mean square error (RMSE) and mean absolute error (MAE) as metrics. The results confirm the superiority of the proposed model.

Cyberbullying Detection Neural Networks using Sentiment Analysis

Jalal Omer Atoum
Department of Computer Science, University of Texas at Dallas, Texas, USA

Abstract- The advances in technical evolution have given rise to a serious problem of cyberbullying. Cyberbullying is the use of electronic communication to bully a person, typically by sending messages of an intimidating or threatening nature. Social networking sites in particular Twitter is becoming a platform for this type of bullying. Machine learning (ML) techniques have been widely used to detect cyberbullying through detecting some language patterns that are exploited by bullies to attack their victims. Sentiment Analysis (SA) of text can also contribute useful features in detecting offensive or abusive content. Deep learning specifically the Convolutional Neural Networks (CNN) has been used to improve the performance of feature extraction during the detection of cyberbullying process. In this research, a SA model is proposed for recognizing cyberbullying tweets in Twitter web-based media. Convolutional Neural Network, Support Vector Machines (SVM) and Naïve Bayes (NB) are utilized in this model as supervised ML classifiers. The aftereffects of the analyses led on this model demonstrated empowering results when a higher n-grams language models are applied on such tweets in comparison with comparable past exploration. Moreover, the results showed that CNN classifiers have outperformed NB and SVM classifiers in several measures.

Data Driven Process Evaluation Simulation Model for Circular Economy Treatment Processes

Dominique Fabio Briechle, Sebastian Lawrenz
Center for Digital Technologies, Clausthal University of Technology, Goslar, Germany

Abstract—The extension of product lifetime is one of the currently most discussed topics regarding the transformation towards a Circular Economy. Primarily electronic devices feature a lifetime development that is highly dependable of the current state of technology evolution. Therefore, the adjustment of product attributes throughout its lifetime becomes more and more essential, and ways to implement those changes of features are required in order to reach the goal of a sustainable environment system. The paper focuses on the problem statement regarding the current state of digitized Circular Economy processes and proposes a simulation layer in order to use artificial intelligence to evaluate the state of a product in use to predict its further treatment after a defect or after the end of its technical compliance with the user requirements.

Sequential Recommendation for Online Consumption by ACFNN

Yahong Lian, Jing Yun, Xiaoming Su, Zhiwei Xu
Institute of Data Science and Application, Inner Mongolia University of Technology, Hohhot, China

Abstract—Online consumption promotes more and more users to purchase stuffs and receive service by Internet. Digging out users' underlying preference chronologically can benefit both customers and servicers. Although several methods have been built on basis of the Markov chain mechanism and Convolutional models, they still insufficiency unveil tangled relationships involved in users' historical behavior. In this paper, we deliver an Attention-based Convolutional Fusion Neural Network (ACFNN) for the sequential recommendation. Specifically, we first conduct items pairwise encoding and then put them into two branches of CNN block, i.e., one is the naive CNN module and another is combined with the attention mechanism. In this way, we can extract more rich representations of sequential features and capture useful item-item correlations. Through comparative experiments, our proposed method achieves better performance by various evaluation metrics.

Multi-Task Deep Neural Networks for Multimodal Personality Trait Prediction

Dena F. Mujtaba, Nihar R. Mahapatra
Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan, USA

Abstract—Organizations, work, and the hiring process are evolving as artificial intelligence (AI) is integrated. A prominent example is video interviews used by large organizations to quickly screen job candidates. The personality traits of job candidates, such as the Big Five characteristics, are predicted using computer vision and affective computing approaches. Past methods have used feature extraction, text, and other multimodal methods to achieve a high prediction accuracy. We build upon past approaches by using multi-task deep neural networks (MTDNN) for predicting personality and the job interview score of individuals. An MTDNN shares lower layers to learn features which apply across outputs, and contains task-specific layers to predict each individual trait, providing an advantage since personality traits are determined with features (e.g., emotion, gestures, and speech) shared between traits. Our model is trained using the CVPR 2017 First Impressions V2 competition dataset, containing 10,000 videos of individuals and their Big Five personality and interview scores. We also use scene, audio, and facial features from the state-of-the-art model from the competition. A 5-fold cross-validation approach is used to evaluate our results. We achieve a prediction accuracy for all traits on par with stateof-the-art individual models, while reducing training time and parameter tuning to a single network.

CSCI-ISHI:
HEALTH INFORMATICS AND MEDICAL SYSTEMS

Applied Blockchain Technology in Saudi Arabia Electronic Health Records

Ashwag Maghraby, Asmaa Numan, Arwa Al Mashi, Asrar Aljuhani, Razan Almehdar, Nada Abdu
Department of Computer Science, Umm Al-Qura University, Mecca, Saudi Arabia

Abstract— One of the major problems facing the healthcare system in the kingdom of Saudi Arabia is difficulty in sharing medical data between hospitals, therefore the patient may suffer when he/she moves from one hospital to another which leads to waste of time, money, and resources. Another problem facing the healthcare system is data loss which could be caused by attackers, corruption, or viruses. Accordingly, this project aims to solve these problems using blockchain technology and creates a peer-to-peer network for healthcare systems in the kingdom Saudi Arabia to store uniform electronic health records. Moreover, this project allows sharing and exchanging electronic health records between hospitals, provides high security and synchronization, and allows access to health information at any time in a faster and easier way.

Smart Task Reminder Model for People with Mild Dementia at Home

Dan O. Oketch, Topside E. Mathonsi
Department of Information Technology, Tshwane University of Technology, Pretoria, South Africa

Abstract - Assistive technologies (AT) for people with mild-tomoderate dementia (PwMMD) has evolved over the years and there is a growing awareness of the benefits to users and the need to identify applicable and effective methods for personalized implementations to cater to individual needs, in this paper, a wrist-worn preprogrammable personalized smart Global System for Mobile Communication (GSM) based task reminder system is proposed, the back-end of the system consists of a computer unit with an internet connection, a bulk SMS subscription service application and a Global System for Mobile Communication (GSM) modem connection, the front-end of the system consists of a user wrist-worn GSM module/receiver for receiving SMS task reminders from the back-end unit via the Global System for Mobile Communication (GSM) network. The aim of this paper is to design and measure the performance through simulation, of a wrist-worn preprogrammable personalized smart Global Sysytem for Mobile Communication (GSM) based task reminder.

Mental Stress Detection with Ensemble Model and Random Cropping using Electrocardiogram

Mingu Kang, Siho Shin, Gengjia Zhang, Jaehyo Jung, Meina Li, Youn Tae Kim
Department of IT Fusion Technology, AI Healthcare Research Center, Chosun University, Dong-gu, Gwangju, Korea;
College of Instrumentation and Electrical Engineering, Jilin University, Jilin, China

Abstract— In this study, we developed a method for classifying mental stress using an ensemble algorithm that combines convolutional neural network (CNN) and k-means clustering. The acquired electrocardiogram (ECG) signal is divided into one-cycle signals, and the data in the time domain are converted to the frequency domain. To improve the accuracy of the mental stress signal classifier, the number of training data was initially augmented using random cropping. Subsequently, an ensemble network based on CNN and k-means clustering was developed to classify the mental stress signals. The confusion matrix and receiver operating characteristic curve were used to evaluate the performance of the model in classifying the signals of mental stress, with an accuracy of 99.2%. With the proposed framework, the mental health management can be improved by classifying the stress in real time.

Data Provenance Management and Application in Clinical Data Spaces

Xiaoyuan Bao, Jingsi Jiang, Kai Zhang
Medical Informatics Center, Peking University Health Science Center, Beijing, China;
Basic Medical Sciences, Peking University Health Science Center, Beijing, China;
Peking University People's Hospital, Peking University Health Science Center, Beijing, China

Abstract—In the era of big data, management of the increasing volume of clinical data has become a research hotspot. To manage the massive data in clinical data centers, and in order to clarify the data source, data iteration and data flow, we propose a data provenance management method based on a three-tier structure, which divides the management of data lineages into a data layer, a semantic layer, and a presentation layer. The method implemented in this paper collects all the commands executed by the operating system, the logs generated by each process, and the physical locations and relevant attributes of all data objects such as data tables, data files, dictionary files, web pages, etc. The semantic layer processes relevant information and generates a data lineage log based on the PROV standard proposed by W3C, and presented it at the presentation layer. Our method records the source, process, and storage location of each process, which enabled process replay. At the same time, our method semantically annotates each data object, which can better describe the source of data items, processing process, etc. Our method enables the replay of the analysis process and the retrospective analysis of the output process. And can be widely used in the management of massive clinical data.

Machine Learning and Big Data Based Clinical EBM Evidence Acquisition and Evaluation

Meng Jin, Jingsi Jiang, Kai Zhang, Xiaoyuan Bao
Medical Informatics Center, Peking University Health Science Center, Beijing, China;
Basic Medical Sciences, Peking University Health Science Center, Beijing, China;
Peking University People's Hospital, Peking University Health Science Center, Beijing, China

Abstract—In the practice of evidence-based medicine, doctors are required to obtain the best medical evidence in real time. However, as far as the medical big data faced by doctors is concerned, it is very difficult to rely on individual doctors to conduct comprehensive inquiry and integration of relevant literature evidence and clinical data. Therefore, in response to the above difficulties, we have initially implemented a platform that is integrated into the information system used by doctors daily to help doctors obtain real-time evidence, sort by quality, and intelligently display; the entire research and development work is mainly aimed at evidence-based medicine practice During the process, we conducted in-depth research on key links such as the best evidence acquisition, evidence quality judgment, and evaluation of the results of evidence application, and realized the computer automation of these links. Under the premise of ensuring the quality of diagnosis and treatment, we tried to improve the acquisition of the best evidence, Evaluation and clinical availability..

Privacy Modelling in Contact Tracing

Marius Molnvik Oye
Department of Computer Science, Norwegian University of Science and Technology (NTNU), Gjovik, Norway

Abstract—Contact tracing is a particularly important part of health care and is often overlooked or forgotten up until right when it is needed the most. With the wave of technological achievements in the last decade, a digital perspective for aid in contact tracing was a natural development from traditional contact tracing. When COVID-19 was categorized as a pandemic, the need for modernized contact tracing solutions became apparent, and highly sought after. Solutions using the Bluetooth protocol and/or Global Positioning System data (GPS) were hastily made available to the public in nations all over the world. These solutions quickly became criticized by privacy experts as being potential tools for tracking. Index Terms—Consumer Electron

Myolink: Emb-Based Inter-Human Wireless Neuroprosthetic Controller

Kathleen Martin, Jimmy Lee, Andrew Kim, Yi Zhao, Shane Walsh, Bahram Shafai
Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts, USA

Abstract - The Myolink team designed and constructed a neuroprosthetic device for inter-human arm control which stimulates muscle contraction in the test subject. This system measures electromyographic (EMG) signals, filters and amplifies the raw signal, then converts the signal from analog to digital. Then the device wirelessly triggers a transcutaneous electrical nerve stimulation (TENS) unit to induce muscle contraction in another person. This system features completely wireless communication between the "controller" and the "controlled" person, as well as multiple muscle targets. When the controller flexes his or her muscle, the corresponding muscle in the controlled person flexes on command. Testing of the Myolink system was conducted in the Electrical and Computer Engineering Capstone Laboratory at Northeastern University in Boston, Massachusetts. An initial experiment was run to verify an existing human-human interface system to confirm placement of electrodes and resulting output motions. Using these observations, the wireless MIMO system was designed, constructed, and tested. The result of the project was a proof-of-concept system that wirelessly induces 2 unique motions generated voluntarily by one person onto a second person via electrical stimulation. Further development of Myolink could be used in applications such as physical therapy, disease recovery tracking, and body movement instruction.

Neural Net Analysis of Peak-Calling Errors in Cut&Tag Data

Micah Lessnick, Holger Mauch
Computer Science, Natural Sciences Collegium, Eckerd College, St. Petersburg, Florida, USA

Abstract— CUT&Tag analysis from multiple EWS/FLI, a fusion protein important in Ewing's Sarcoma, cell lines have a potential to return false positives as a result of DNA folding. The current approach to determine which peaks are false and which are true involves aligning DNA and categorizing peaks based on overlapping locations. This project uses a neural network to build a model based on the features of each peak and use that as the basis for identifying false positives from true peaks. After training a densely-connected neural net model, the best-performing model was able to identify false peaks from true peaks with an F1 score of 0.82 on the training set and 0.67 on an additional testing cell line.

Middleware to Integrate Patient Data from Heterogeneous Distributed Databases and its Efficacy

Subrata Kumar Das, Mohammad Zahidur Rahman
Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh

Abstract—The health organizations store the patient data in different repositories and scattered in diverse locations. In the healthcare domain, the problem is that each hospital or even each department under a hospital maintains its own database having various data models (SQL, NoSQL, etc.). In this situation, existing or new applications require to grant healthcare actors to locate and share patient data from those pre-existing distributed databases (DDBs) remotely for the needs of patient quality treatment, daily operations of the health centers. However, data integration from distributed data sources is raising concern for data model variability. Therefore, it is significant to identify that how much an application like middleware is efficient to reconstruct and share patient data remotely from heterogeneous DDBs over the networks. The health organizations could also require to ensure whether their existing database model performs well or should replace by another one. So, this paper aims to design a system using different databases consisting of distinct data structures and an algorithm for middleware to integrate data from them with testing the system performance. The experimental results of this research work show that the patient data could be shared from various distributed data sources efficiently. Therefore, the study could direct the healthcare organizations for sharing patient data from heterogeneous distributed databases without replacing the existing data model.

Development of a Causal Model to Study the Disparate Effects of COVID-19 on Minorities

Misagh Faezipour, Miad Faezipour, Brianna Bauman
Department of Engineering Technology, Middle Tennessee State University, Tennessee, USA;
School of Engineering Technology, Electrical & Computer Engineering Technology,
Purdue University, West Lafayette, Indiana, USA;
Department of Engineering Technology, Middle Tennessee State University, Tennessee, USA

Abstract—This paper proposes a systems engineering perspective to analyze the causes of COVID-19 health disparities impact and interventions to minimize the impact on minorities. The impact of the novel coronavirus has shown to be more intense on minorities. The percentage of COVID-19 case count and fatality rate for minorities is much higher than that of the general population, showing that they are more significantly affected than others. Many different factors influence this impact, ranging from economic to cultural. In this paper, these factors are shown to be connected through a causal model analyzing the effects of each factor, after which, potential interventions are suggested. Many factors are identified, such as high employment in the service industry or lower likelihood to have insurance. From this, a causal model is created showing the impact of each factor. Using this causal model, one can identify the high-impact factors causing a disparate impact as well as suggest possible interventions including making testing and treatment more accessible, reducing healthcare bias, and improving healthcare for immigrants.

Surgical Skill Training and Evaluation for a Peg Transfer Task in a Three Camera Based-Laparoscopic Box-Trainer System

Fatemeh Rashidi Fathabadi, Janos L. Grantner, Saad A. Shebrain, Ikhlas Abdel-Qader
Electrical & Computer Engineering, Western Michigan University, Michigan, USA;
Surgery, Homer Stryker M.D. School of Medicine, Western Michigan University, Michigan, USA

Abstract— In laparoscopic surgery, surgeons should acquire additional skills before carrying out real operative procedures. The manual skills component of the Fundamentals of Laparoscopic Surgery exam is essential to measure the trainees' technical skills. The peg transfer task is a hands-on exam in the FLS program. In this paper, a multi-object detection method is proposed to improve the performance of a laparoscopic boxtrainer-based skill assessment system from top, side, and front cameras. Based on experimental results, the trained model could identify each instrument at a high score of fidelity and the trainvalidation total loss for the SSD ResNet50 v1 FPN was about 0.06. In addition, this method could correctly identify the peg transfer time, the move, the carry and dropped states of each object from top, side, and front cameras. This improved intelligent laparoscopic surgical box-trainer system helps in enhancing surgery residents' laparoscopic skills. This project is a collaborative research effort between the Department of Electrical and Computer Engineering and the Department of Surgery, at Western Michigan University.

Modeling Clinical Practice Guidelines for Interactive Decision Support Exemplified by Primary Myelofibrosis and Immune Thrombocytopenia

Patrick Philipp, Louisa Hempel, Dirk Hempel, Jurgen Beyerer
Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhofer Center for Machine Learning, Karlsruhe, Germany; Vision and Fusion Laboratory IES, Karlsruhe Institute of Technology KIT, Karlsruhe, Germany;
Steinbeis-Transfer-Institut Klinische Hamatoonkologie, Donauworth, Germany

Abstract—Clinical Practice Guidelines (CPGs) contain expert knowledge on the diagnosis and treatment of diseases. They can be regarded as state of the art and standardized procedures that have been established by consensus of the clinical expert community. In this work, we show how CPGs can be formalized by activities of the Unified Modeling Language (UML), and can subsequently be translated into PROforma models. UML activities allow for a comprehensible representation of the underlying process, whereas PROforma models can be directly executed in a dialog system and support the practitioner during the diagnosis or treatment process. In this work, we expand our approach from [1] to include more complex diseases like Primary Myelofribrosis (PMF) and Immune Thrombocytopenia (ITP) and show the applicability for exemplary patients.

Iterative Ensemble Transductive Learning for Microscopy Image Analysis

Tian Shi, Longshi Wu, Changhong Zhong, Ruixuan Wang, Hongmei Liu
Sun Yat-sen University, Guangzhou, China;
GuangDong Province Key Laboratory of Information Security Technology, Guangzhou, China

Abstract—In automatic histopathology and microscopy image analysis, due to high patient-level variability, the model trained based on the images from a set of patients may not perform well on the images from a another set of patients. To overcome this issue, motivated by transductive learning and ensemble learning, we propose an iterative framework to train ensemble transductive models with pseudo-labels. In each iteration, a number of single models are first trained by combining the training set with part of randomly selected test data which have high prediction confidence, and then ensembled to predict the labels of test set for the next iteration. In this way, the latent information in test set would be exposed to the model such that the model can directly learn from the test data to be analyzed. Experimental evaluation on two datasets, white blood cancer microscopic images and breast histopathology images, shows that the proposed approach significantly outperforms the traditional ensemble models.

A Novel Event-Related Desynchonization/Synchronization with Gamma Peak EEG Model for Motor State Identification

Ravi Suppiah, Anurag Sharma, Noori Kim, Khalid Abidi
Electrical and Electronic Engineering, Newcastle University in Singapore, Singapore

Abstract - Electroencephalography (EEG) is one of the imperative physiological signals for therapeutic and rehabilitative robotic control. The EEG features in time-domain, frequency-domain, and time-frequency have been extensively studied, especially differentiating between right and left movements of hands and legs. However, there are limitations in addressing real-world practical implementation due to multiple channels and computationally intensive and complex algorithms. This research proposes a novel model that can provide a high-level accuracy using only a single channel with straightforward signal processing algorithms using Event-Related Desynchronisa-tion (ERD) and EventRelated Synchronisation (ERS) signals. We demonstrate to distinguish between the arm's 'Rest' and 'Active' states by using a single-channel EEG signal. It is a practical model that can easily extend to any embedded device for a portable realworld application.

What Do I Want? Supporting the Incorporation of Individual Patient Preferences for Decision Support in Breast Cancer Therapy

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

Abstract—With the rise of personalized medicine, the number of individualized treatment options and related decisions is increasing tremendously. Thereby, the acquisition, incorporation and representation of the patient's individual preferences in upcoming, modern, AI-based medical decision support systems play a decisive role. E.g., for patients with advanced breast cancer, there are various therapeutic options associated with different outcomes to choose from. In our contribution we show a first approach to model Preference Elicitation (PE) via card sorting using a utility function. Based on this, we present further ideas for extending and improving the approach.

Motor State Classification based on Electromyography (EMG) Signals using Wavelet Entropy and Neural Networks

Ravi Suppiah, Khalid Abidi, Noori Kim, Anurag Sharma
Electrical and Electronic Engineering, Newcastle University in Singapore, Singapore

Abstract - Neuromuscular system works based on communication between nerves and muscles. Capturing neuromuscular activities is employed in many application domains, ranging from remote-robotic control, rehabilitation, remote surgery, etc. The ability to accurately decode the intention determines the reliability of the application. This paper proposes a motor state classification technique based on Electromyography (EMG) signals. Current research in the general domain of EMG signal classification has generated good results. However, the existing techniques require high computing resources or are nonreal-time in nature, making them impractical for real-world application. In this paper, the authors propose the use of wavelet entropy and neural networks for classification of four major hand movements, forward, reverse, raise and lower. The results ranging from 91.9% to 93.5%show promise in the extension of the proposal for further classification.

Impact of Patch Extraction Variables on Histopathological Imagery Classification Using Convolution Neural Networks

Willmer R. Quinones, Murtaza Ashraf, Mun Yong Yi
Graduate School of Knowledge Service Engineering, KAIST, Daejeon, South Korea

Abstract—Pathologists visually examine cell morphology by observing the biopsy slides under a microscope through different magnifying factors. This process is time-consuming and error-prone. In this regard, computer-aided whole slide image analysis is necessary to help pathologists reduce time effort and human error. With the recent advances in deep learning for computer vision, convolutional neural networks (ConvNets) started gaining attention in the medical domain and have shown significant progress in whole slide image classification. Existing deep learning approaches works by extracting small patches from a whole slide image (WSI) and using them for ConvNet training. However, it is unknown what effects the patch extraction variables of the size of the patches and the magnifying factor of the WSI have on the performance of the ConvNet. Therefore, we construct several datasets by extracting patches from stomach histopathological imagery. Generated datasets contain different sizes of the patches and the magnifying factors of the WSI. Densely Connected Convolutional Network (DenseNet) assessed to classify dysplasia, malignant and benign patches. The impact of patch extraction variables was observed using precision and recall. This study shines light on why these factors would affect the model performance concerning data representation and provides a guideline for histopathological imagery data extraction methods.

Toward Intelligent Respite-Care Dyad Monitoring Using IoT

Shakour Abuzneid, Miad Faezipour
Department of Computer Science, St. Cloud State University, St. Cloud, MN, USA;
School of Engineering Technology, Electrical & Computer Engineering Technology,
Purdue University, West Lafayette, Indiana, USA

Abstract—The relationship between care recipients such as the elderly, day-care children, critically ill people, and persons with dementia (PWD) on one side, and the caregivers on the other side, is complex. This relationship could cause agitation episodes continuously. Caregiving is challenging because of the high workloads and human-related conditions such as fatigue, lack of sleep, and other health situations. This relationship could go through mutual hardship, which only needs to be predicted, detected, prevented, and normalized. It was proven that it is hard to predict such relationships with pre-screening processes because it is evolving. However, we should not wait until abuse to the care recipients accumulates to unacceptable levels. With the advent of technology, including sensing, monitoring, Internet of Things (IoT), cloud computing, and artificial intelligence (AI), continuous, uninterrupted intelligent systems can be developed to reveal the relationship between the dyads without exposing privacy. The system under development is an IoT intelligent system that derives data about the dyad's relationship and exports data to the cloud to apply AI techniques in real-time. The output will be in the form of reports and alarms shared with proper management, indicating

highly critical evidence of abuse. It also provides an evaluation of caregiver performance. In Phase I, the system will be tested in an experiential environment. After adjusting the system based on the results of Phase I, we will evaluate the system in an actual elderly care facility in Phase II.

A Review of Publicly Patient-Centered Alzheimer's Disease Datasets

Rui Wang, Chwan-Li Shen, Victor Sheng, Heng Wu, Zackery Toler, Samuel Temesgen
Department of Pathology, Texas Tech University Health Sciences Center, Lubbock, Texas, USA;
Department of Computer Science, Texas Tech University, Lubbock, Texas, USA;
Department of Mathematics and Computer and Science, West Virginia State University, Institute, West Virginia, USA

Abstract—Alzheimer's disease (AD), the major cause of dementia, is becoming a global health issue. This review includes major publicly available patient-centered AD datasets, covering three main categories of data – clinical, imaging, and genetic, to facilitate researchers to pinpoint suitable dataset(s) for their studies. It overviews these datasets with their associated studies, data storage locations, and most importantly, details of the data such as cognitive exam results, biomarker measurements, neurological images, genotyping and expression data. They provide an informative and useful portal for AD or Machine Learning researchers and can help them pinpoint the desired datasets suitable for their studies.

A User Self-Design Collaborative Web System for Self-Directed Health Learning: A Covid-19 Awareness Project

Alireza Ebrahimi
State University of New York at Old Westbury, Old Westbury, New York, USA

Abstract - The aim of the system is to direct the user to create their own network system collaboratively for a case of Covid-19 health throughout a case study in a Web System and E-Commerce course. The framework of directed health learning is based on 7x2C content knowledge established criteria to assess the effectiveness and efficiency of the system. The self-design frame framework is plan oriented and based on the concept of plan and plans integrations and special relationships. The user is directed to break the system into plans and the design is to self-guide the user to build a system to comprehend, combat, coexist, cope and traces COVID-19 with four layers of diagnostics, simulation, and pattern matching database. With collaboration of users' systems, a fact from one user as output can be transferred to another user as an input in a circulation forming a general fact. Consequently, the transfer of learning from one system flows into another system resulting in a pattern to be found. Based on the pattern an algorithm will formulate to tackle a solution to COVID-19. The implication of this study will be a guideline for others to initiate their own participant's system to find a pattern and formulate an algorithm for the pandemic. The idea of self-design and self-directed learning can be transfer to other fields of study covid-19 health. At present time, a parallel case study of goods and services on farming of Sunchoke plant has been directed with three plan themes of grow, eat, and heal.

Sensor-based System to Algorithmically Recognize Gait Disturbances Common to Vascular Dementia Onset

Arshia Khan, Janna Madden
Department of Computer Science, University of Minnesota Duluth, Duluth, Minnesota, USA

Abstract - Early detection of Vascular Dementia onset is challenging in a clinical setting, as the earliest signs often present in physiological aspects. One major aspect of this is gait patterns. This project utilizes force-sensing platforms, motion capture, and EMG sensors to unobtrusively collect biometric data from an individual's walking pattern. Following data collection, a series of algorithms computes statistics off the gait cycles. In addition to previously validated biometric indicators of vascular dementia, including stride length, time in stride and swing phases of gait, time in dual leg vs single leg support, this system also examines metrics surrounding balance, lateral movement, and fine-grained gait analysis during critical transition periods of gait, when weight is transferred from one leg to the other. The proposed system provides a tool for which gait can be analyzed and compared over a long period of time. Such a system opens opportunity to increased personalization in health monitoring and disease diagnosis and provides an avenue to increase patient-centricity of medical care.

Estimating the Trend of COVID-19 in Iran Before and After the Start of Vaccination

Alireza Safarzadeh, Mohammad Javad Heidarian, Reza Sharifanfar, Mohammad Reza Jamali, Sina Forqandoost Haqiqi
School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran;
Pulseware Co., Tehran, Iran;
Lanvera LTD., Frisco, Texas, USA

Abstract—The newly emerging coronavirus 2 (SARS-CoV-2), known as COVID-19, is severe acute respiratory syndrome. On 11th March 2020 this virus has been recognized as a pandemic by the World Health Organization (WHO). There are many unknowns about this virus to date. The prime objective is to control from spreading worldwide. This paper used six mathematical model for predicting COVID-19 virus behavior and effects of vaccination on controlling the spreading the virus in five countries with special focus on Iran. The modeling studies are conducted for Germany, USA, Turkey, and Israel. In the following sections, each country is studied and finally the trend in Iran analyzed and compared with other countries. The results showed that, after vaccination of more than 50% of the population, the mortality rate due to COVID-19 decreased by more than 70%. Also results showed that High-paced vaccination led to a decrease of about 10.35 million fewer COVID19 cases and about 240 thousand fewer death cases in these countries.

Decision Making on School Closure Based on Correlation Analysis During COVID-19 Pandemic

Iyke Maduako, Kwon Ryul Cha
Office of Global Innovation, UNICEF, New York, USA;
Cumberland Valley HS, Mechanicsburg, PA, USA

Abstract— A significant outcome of the COVID-19 pandemic has been the hyper dependency on digital tools. As schools and, at many times, businesses have converted online, the use of digital tools has increased rapidly. Since the situation of COVID-19 fluctuates, national and state level regulations have caused schools to switch back and forth between traditional and online formats. When deciding between the two, governments should factor in the effects on mental health to prevent increased rates in depression. Our research aims to analyze the correlation between depression rates and the use of digital tools to implement an autoregressive model to predict the best times to switch between online and traditional teaching formats. Based on our findings, the auto regressive model will be implemented to predict when depression rates would peak after school closures.

An Early Warning System for Patients in Emergency Department based on Machine Learning

Ying-Feng Hsu
Osaka University, Japan

Abstract—In this paper, we present an early warning system for patients in the emergency department. Our proposed system includes data processing steps that transform raw clinical data streams into useful information that facilitate clinical decision making for the early warning. We tested the proposed approach in a medical monitoring system, which takes physiologic data and predicts in which clinical setting the data is most likely to be seen. To demonstrate high utility of our approach, we conducted a set of experiments on the clinical data of 1,176 patients.

CSCI-ISCI:
COMPUTATIONAL INTELLIGENCE

Towards 'Serendipity Engineering for Seductive Hypermedia' and 'User Analysis using Socialnomics': The Role of Ecological Cognition

Jonathan Bishop, Mark M. H. Goode
Congress of Researchers and Organisations for Cybercommunity, E-Learning and Socialnomics, Berkeley, England, UK;
Cardiff Business School, Cardiff University, Cardiff, Wales, UK

Abstract— This paper investigates the long-established ecological cognition framework and updates it to better account for the advances in computational science computational intelligence. To do this, two concepts are explored. The first, 'serendipity engineering for seductive hypermedia,' looks at how to design information systems to account for the pleasant occurrences that happen in offline environments studied by those in sales and marketing where beneficial outcomes often occur by chance encounters. The second, 'user analysis using socialnomics' looks at how a parametric user model based on the ecological framework can be used to understand users of information systems from the point of view of supporting a digital economy of users. A number of additional equations are developed using socialnomics that can be applied to digital transformation based on the parametric user model, including to calculate probability of seduction and probability of serendipity in an information system. The parametric model presented has great applicability for information and communications technology solution providers.

Abstractive Text Summarization via Stacked LSTM

Ireddy Siddhartha, Huixin Zhan, Victor S. Sheng
Department of Computer Science, Texas Tech University, Lubbock, Texas, USA

Abstract—In the past, there have been many models proposed for text summarization via sequence to sequence training (seq2seq), attention mechanism, and transformers. Although these methods achieve an advance regarding the performance, these models fail to create a more complex feature representation of the current input and consequently gain inferior performance for modeling the long staggered sentences and modeling the complex inter-sentence dependencies. In order to address this issue, we utilize a more complex feature representation for summarization via stacked LSTM. In this case, the main reason for stacking LSTM is to allow for greater model complexity. For a simple encoder, we stack layers to create a hierarchical feature representation with attention. We generate the text summaries for any test text in terms of predicting the target sequence. With the proposed method, we achieve a better performance compared to the existing state-of-the-art phrase-based system on the task of text summarization on gigaword dataset. Furthermore, Experimental results on this dataset show that our framework performs well in terms of various ROUGE scores.

Accordion Autoencoders (A^2E) for Generative Classification with Low Complexity Network

Kevyn Kelso, Byeong Kil Lee
Department of Electrical and Computer Engineering, University of Colorado, Colorado Springs (UCCS), Colorado, USA

Abstract— Deep learning technologies are popularly used in many areas including recognition, identification, anomaly detection, classification, etc. Large network complexities requiring more computational resources and response time is a popular challenge. Generative models used in modern applications have lower dimensional latent spaces which can be manipulated to change features of the output. Recently, autoencoders have been used in dimension reduction, data reconstruction, etc. One of the disadvantages is the network size in which the architecture of the autoencoder has a double workload due to the encoding and decoding processes. In this paper, we explore a scheme to reduce the network complexity of the autoencoder. The motivation behind exploring different autoencoder architectures lies in their practical uses for the applications such as anomaly detection, classification, and their usages in generative models. Additionally, a deeply understanding of the mechanism to improve dimensionality reduction may lead to a better understanding of how the human brain finds meanings in data. We come up with an Accordion Autoencoder (A2E) architecture, as an effective performing solution to anomaly detection and classification problems, which is not necessarily a solution to dimensionality reduction, but rather a performance improvement for the problems using several sets of lower-dimensional space to generate more meaningful features of the data. Based on our experiments, the proposed solution provides the network size reduction (85.1%~94.5%) with maintaining the accuracy (4.9%~13.6% accuracy drop) in fraud detection and MNIST classification.

Kabuki Explanation System Based on User's Knowledge and Interests

Miku Kawai, Jumpei Ono, Takashi Ogata
Faculty of Liberal Arts, The Open University of Japan, Wakaba, Mihama, Chiba, Japan;
Faculty of Software and Information Technology, Aomori University, Seishincho, Edogawa, Tokyo, Japan;
Faculty of Software and Information Science, Iwate Prefectural University, Sugo, Takizawa, Iwate, Japan

Abstract—This paper introduces the generation of explanations for a stage-performing structural simulation system that we have developed. First, we classify the collected explanations. Thereafter, we prototype a mechanism in which the system automatically determines the content and method of an explanation based on arbitrary parameters. Through this trial, we will examine effective explanation methods for story generation.

Selecting Optimal Portfolio in Generalized Feed Forward Networks and Support Vector Machines Hybrids

N. Loukeris, I. Eleftheriadis, G. Gikas
University of West Attica, Athens, Greece;
Athens University of Economics and Business, Athens, Greec;
University of Macedonia, Thessaloniki, Greece

Abstract—We evaluate the performance of 70 Generalized Feed Forward and 14 Support Vector Machines models of plain and hybrid form to define the optimal classifier in portfolio selection.

Sublinear Evaluation of Complex Networks for Extensive Exploration of Configurations for Critical Scenarios and Decision Making

Alexandre Delbem, Antonio Mauro Saraiva, Joao Bosco A London Jr, Rodrigo Fanucchi
Institute of Mathemathics and Computer Science, University of Sao Paulo, Sao Carlos, Brazil;
Escola Politecnica, University of Sao Paulo, Sao Paulo, Brazil;
Sao Carlos School of Engineering, University of Sao Paulo, Sao Paulo, Brazil;
Department of Maintenance and Distribution Operation, COPEL, Curitiba, Brazil

Abstract—Artificial Intelligence in Complex Networks has contributed to several relevant fields involving energy systems, computer networks, environment, agriculture, health, and social organizations. However, investigations concerning multiple heterogeneous networks have been less frequent. Mixed systems usually require fine-grained data to retain a sufficient amount of details from each network. This type of modeling may enable the investigation of emerging behaviors or synergies. For example, decision making may involve the search for an improved network configuration (involving coarse and fine modifications on devices, procedures, and settings) with the lowest possible cost to soon mitigate effects from climate changes or other types of ”attacks” (from economic crises, calamities, and recent pandemics). The generation of robust configurations for heterogeneous networks involves some challenges, pointed out in this paper. The efficient calculus of load flows has been one of the main challenges. To overcome it, we propose a load flow algorithm with sublinear time complexity for the construction and evaluation of several configurations. The new algorithm scales well and can deal with the nonlinear dynamics that evaluations of entire sets of finegrained modeled networks may involve.

Knowledge Inference and Knowledge Completion Methods using Neuro-Symbolic Inductive Rules

Won-Chul Shin, Hyun-Kyu Park, Young-Tack Park
Computer Science Department, Soongsil University, Seoul, South Korea

Abstract— In knowledge graph completion, a symbolic reasoning method establishes a human readable rule by analyzing an imperfect knowledge graph and infers knowledge omitted by an inference engine. However, the entire rules cannot be defined based on a large-scale knowledge graph. This study proposes a method, based on a knowledge graph, that can facilitate end-to-end learning and induce rules without several processing steps that require direct human involvement. The proposed method combines the concept of unification used in symbolic reasoning and deep learning for training vectors expressing symbols. It trains the vectors expressing relations of rule schemas defined to induce rules based on a given knowledge graph. Furthermore, the performance of the proposed method is evaluated against neural theorem prover and the greedy neural theorem prover, which are recently developed neurosymbolic models, based on four benchmark datasets. The experimental results verify that the proposed method induces more significant rules in less training time. Furthermore, this study conducted an experiment on knowledge graph completion, implemented by an inference engine. Based on the experiment results, it was confirmed that the rules induced by the proposed model can indeed effectively complete missing knowledge.

Hybrid Meta-heuristics Approach for Solving Supply Chain Network Model under Disruption Risk

Chuluunsukh Anudari, YoungSu Yun, Mitsuo Gen
International Cooperation Team, Chosun University, Gwangju, Korea;
School of Business Administration, Chosun University, Gwangju, Korea;
Fuzzy Logic Systems Institute and Tokyo University of Science, Japan

Abstract—A supply chain network (SCN) model which considers facility and route disruptions simultaneously is proposed in this paper. Since most of conventional literatures have focused either on facility disruption solely or on route disruption solely, the simultaneous consideration of facility and route disruptions can improve the flexibility of the implementation in the SCN model. The SCN model under the disruptions is represented as a mathematical formulation and a hybrid meta-heuristics (GA-VNS) approach which combines genetic algorithm (GA) with variable neighborhood search (VNS) is used for the mathematical formulation. In numerical experiment, two scaled SCN models are used for comparing the performance of the GA-VNS approach with those of some conventional meta-heuristics approaches. Experimental results prove that the GA-VNS approach is more robust than conventional meta-heuristics approaches, and the flexibility of the SCN model under the disruptions are also improved.

Hybridizing Teaching-Learning Based Optimization with GA and PSO: Case Study of Supply Chain Network Model

Mitsuo Gen, Chuluunsukh Anudari, YoungSu Yun
Fuzzy Logic Systems Institute and Tokyo University of Science, Japan;
International Cooperation Team, Chosun University, Gwangju, Korea;
School of Business Administration, Chosun University, Gwangju, Korea

Abstract—Design and optimization of logistics and supply chain management (SCM) network is a very important issue, which plans, implements and controls the efficient, effective forward and reverses flows and storage of goods, services and related information between the point of origin and the point of consumption to meet customers' requirements. The structure of each stage of the supply chain is completed by a set of arcs connected to the nodes and representing the flow of goods: each arc has a weight proportional to the cost incurred in between the two nodes it connects together. In this paper, we first formulate the nonlinear 0-1 integer programming model of the SC network consisting of suppliers, manufacturers, DCs, and customers. Secondary we propose a hybridized teaching-learning based optimization (TLBO) with genetic algorithm (GA) and particle swarm optimization (PSO) combined Cauchy distribution to solve the SC network models. Lastly, we demonstrate the proposed GA+PSO+TLBO method outperformed GA, TLBO, GA+TLBO under the same computational environment for two numerical experiments of the several scaled SC network models.

Optimal Balance of Privacy and Utility with Differential Privacy Deep Learning Frameworks

Olivera Kotevska, Folami Alamudun, Christopher Stanley
Computer Science and Mathematics, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA;
Computational Sciences and Engineering, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

Abstract—As the number of online services has increased, the amount of sensitive data being recorded is rising. Simultaneously, the decision-making process has improved by using the vast amounts of data, where machine learning has transformed entire industries. This paper addresses the development of optimal private deep neural networks and discusses the challenges associated with this task. We focus on differential privacy implementations and finding the optimal balance between accuracy and privacy, benefits and limitations of existing libraries, and challenges of applying private machine learning models in practical applications. Our analysis shows that learning rate, and privacy budget are the key factors that impact the results, and we discuss options for these settings.

Ant Colony Optimization for Balanced Multiple Traveling Salesmen Problem

Sun Bing, Wang Chuan, Yang Qiang, Liu Weili, Yu Weijie
College of Computer and Information, Henan Normal University, Henan, China;
College of Software, Henan Normal University, Henan, China;
College of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, China;
School of Computer Science, Guangdong Polytechnic Normal University, Guangdong, China;
School of Information Management, Sun Yat-sen University, Guangdong, China

Abstract—The balanced multiple traveling salesmen problem (BMTSP) is a popular combinatorial optimization problem of real world, which has two objectives: minimizing the total path length of salesmen and minimizing the longest path of salesmen. In this paper, a balanced ant colony optimization (BACO) algorithm is proposed to solve balanced MTSP. To be specific, firstly, BACO solves the MTSP by optimizing the ant groups, each ant stands for a salesman in a group. Secondly, a selection mechanism is proposed to optimize the balance of paths in the path selection process. This mechanism has four modes, namely Random Selection (RS), Shortest Best Selection (SBS), Future Based Balance Selection (FBBS) and Future Based Short Selection (FBSS). These modes can regulate the path selection process of salesmen and optimize the solution. Additionally, we introduce the 2-opt operation to optimize the path of each ant. Finally, the proposed BACO is compared with the improved genetic algorithm (GA) on TSPLIB benchmark instances with four test sets. The comparison results have vilified the effectiveness of the BACO.

Application of an Input-Output Pairings Selection Methodology to Control Multivariable Systems based on Multi-Objective Optimization

Victor Huilcapi, Xavier Blasco, Juan M. Herrero, Alberto Pajares, Gilberto Reynoso-Meza
Facultad de Ingenierias, Universidad Politecnica Salesiana, Guayaquil, Ecuador;
Instituto Universitario de Automatica e Informatica Industrial, Universitat Politecnica de Valencis, Valencia, Spain;
Programa de Pos-Graduacao, em Engenharia de Producao e Sistemas (PPGEPS),
Pontificia Universidade Catolica do Parana (PUCPR), Curitiba, Brazil

Abstract—This paper shows a methodological approach to selecting input-output pairings and tuning controllers for multivariable systems based on multi-objective optimization. The methodology is applied to a nonlinear multivariable system called a boiler-turbine unit (electrical power generation system). Two scenarios and two design concepts (input-output pairings and controllers) were defined, each with six design objectives. It was shown in detail that one design concept is better than the other in each scenario for tracking a reference. It is important to mention that the proposed methodology establishes a unified framework to select suitable loop pairings and tune multivariable controllers defining what in this paper is known as a design concept. The methodology establishes a global comparison framework of each design concept so that a designer according to his/her preferences selects an optimal solution to control the system.

Splatter: An Efficient Sparse Image Convolution for Deep Neural Networks

Tristan Lee, Byeong Kil Lee
Department of Electrical and Computer Engineering, The University of Colorado, Colorado Springs (UCCS), Colorado, USA

Abstract—Deep neural network (DNN) based approaches, such as deep convolutional neural network (CNN), have achieved highly accurate results in many fields (e.g., computer vision, etc.) at the cost of a huge number of parameters and high computational workloads. The parameters require large memory capacity and memory access time which cause a migration problem to embedded devices. Pruning techniques can reduce the DNN complexity, but it brings sparsity in the matrix which causes computational inefficiency and performance loss. The reasons for the inefficiency include the reduced data reuse opportunities, waste of memory bandwidth, and computational irregularity. Applying sparse matrix formats can help to reduce inefficiency with a regular computational pattern of the sparse matrix (e.g., Compressed Sparse Row), but it has a limitation to improve the efficiency in data reuse and memory bandwidth. In this paper, we propose the Splatter which is an efficient sparse image convolution technique for DNN. In the convolution sweep, non-zero input data is multiplied by each kernel element and the outcomes will be accumulated into the output. We focus on reducing memory access and increasing data reuse. The performance of our proposed technique is compared to naïve convolution method with a dense matrix format and CSR format as well. Our experimental results with sparse images, which are 50%~90% sparsity, show that the Spatter can improve the execution time of image convolution by 25%~81% with dense matrix, and 49%~90% with a CSR matrix format. Additionally, an average reduction in input data accesses of 97% is observed using the proposed convolution method.

Zero-Shot Cross-Lingual Transfer in Legal Domain Using Transformer Models

Zein Shaheen, Gerhard Wohlgenannt, Dmitry Mouromtsev
ITMO University - Russia, Saint Petersburg, Russia

Abstract - Zero-shot cross-lingual transfer is an important feature in modern NLP models and architectures to support low-resource languages. In this work, We study zero-shot cross-lingual transfer from English to French and German under Multi-Label Text Classification, where we train a classifier using English training set, and we test using French and German test sets. We extend EURLEX57K dataset, the English dataset for topic classification of legal documents, with French and German official translation. We investigate the effect of using some training techniques, namely Gradual Unfreezing and Language Model finetuning, on the quality of zero-shot cross-lingual transfer. We find that Language model finetuning of multi-lingual pre-trained model (M-DistilBERT, M-BERT) leads to 32.0-34.94%, 76.15-87.54% relative improvement on French and German test sets correspondingly. Also, Gradual unfreezing of pre-trained model's layers during training results in relative improvement of 38-45% for French and 58-70% for German. Compared to training a model in Joint Training scheme using English, French and German training sets, zero-shot BERT-based classification model reaches 86% of the performance achieved by jointly-trained BERT-based classification model.

CSCI-ISCC:
CLOUD COMPUTING AND DATA CENTERS

A Review of Vehicular Micro Clouds

Abhishek Phadke, F. Antonio Medrano, Stanislav Ustymenko
Department of Computer Science, Texas A&M University, Corpus Christi, Texas, USA

Abstract - Data-intensive applications that dominate recent computing have increased the demand for flexible, secure, powerful, and cost-effective provision of computing resources. This article proposes using an extension to cloud computing techniques, which currently enable on-demand delivery of processing, storage, and platform resources over the Internet. Cloud services are traditionally provided by central data centers, usually located far away from end-users. Physical distance and uncertain communication infrastructure can potentially degrade the quality of available services. To combat these issues, earlier studies introduced Edge computing and local cloudlet architectures. The emerging idea of micro-clouds placed closer to the user has been observed to overcome these issues. Vehicular micro-cloud (VMC) deployments take advantage of the extensive computing power and communication equipment in modern vehicles to deploy mobile cloud solutions. In this study, we provide the review of literature on vehicular micro-clouds, identify research trends, discuss algorithms for VMC formation, and classify the proposed architectures based on their deployment strategy. We also present the novel idea of employing incentivization protocols aimed to create bigger and more stable VMC instances.

An Algorithm to Enhance Data Integrity in Cloud Computing

Moloko P. Mothlabeng, Topside E. Mathonsi, Tonderai Muchenje, Deon P. Du Plessis
Department of information Technology, Tshwane University of Technology, Pretoria, South Africa

Abstract—Cloud computing is a standard for providing information technology services over the Internet such as hardware, software, networking, and storage, which can be accessed from anywhere in the world at any time. However, due to the nature of cloud computing, it is difficult to store private data on cloud servers. Data is sent to a remote cloud server, which is insecure and untrustworthy. That means data might be altered intentionally by unauthorized users or accidentally by the service provider causing data loss or data modification. As a result, it's become critical to protect data from hackers to maintain its management procedures, confidentiality, privacy, and integrity. This paper proposes EDIE algorithm to enhance data integrity in cloud computing. The proposed algorithm is a hybrid-based approach designed by fitting in MTD (Moving Target defense) to the New Lightweight Cryptographic algorithm. The EDIE algorithm will use MTD to dynamically change network configurations and frequently confuse the attackers which in turn minimizes the level of attacks. Therefore, leading to an enhanced data integrity. The algorithm uses a 16-byte (128 bit) block cipher and requires 16-byte (128 bit) keys for encryption. The secret key's length and the number of turns is also adjustable. In comparison to other commonly used cryptographic schemes in cloud computing, the proposed algorithm is expected to have lower attack level and improved execution time resulting in enhanced data integrity in cloud computing.

MOBDroid2: An Improved Feature Selection Method for Detecting Malicious Applications in a Mobile Cloud Computing Environment

Noah Oghenefego Ogwara, Krassie Petrova, Mee Loong (Bobby) Yang
School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand

Abstract— This paper presents an ensemble machine learning (ML) based system for the detection of malicious applications in the Mobile Cloud Computing (MCC) Environment. The proposed system named MOBDroid2 applies a static feature analysis approach using the permissions and intents demanded by Android apps. The experiments conducted showed that the proposed system was able to effectively detect malicious and benign apps, achieving a classification accuracy rate of 98.16%, a precision rate of 98.95%, a recall rate of 98.20%, and a false alarm rate of 1.85%. The results obtained in our experiment compared well with other results reported in extant literature.

Utilizing HAPS Deployed Data Centers in Offloading Cloud Workloads

Tolga Ovatman, Muhammed Enes Tirnakci, Ahmet Yilmaz
Department of Computer Engineering, Istanbul Technical University, Istanbul, Turkey

Abstract—Recent advances in communication technology has begun to push the data centers towards high altitude platform stations (HAPS) due to their advantages in communication coverage, mobility and cooling. There are plenty of studies that focus on the communication aspects of such devices but the issues regarding deploying micro data centers on HAPS is rarely studied. In the future, there is a potential for a substantial amount of processing power to be offload to such stations to back up terrestrial stations in fog computing scenarios. In this study we have analyzed cloud workload handling and power consumption efficiencies of two different HAPS deployment scenarios. In our simulations, performed with CloudSim Plus, we have experimented with scenarios with varying number of lower and higher altitude HAPS and terrestrial base stations. In our experiments we have identified the trade-offs and cases to provide better performance in handling cloud workloads.

Persisting Trust in Emerging Hybrid-Clouds and 6G Systems

Muhammed Akif Agca, Djamel Khadraoui, Mehmet Naci Efe
TOBB ETU Computer Engineering, Ankara, Turkey;
Luxembourg Institute of Science and Technology / LIST, Esch-Sur-Alzette, Luxembourg;
Ekol Group, Istanbul, Turkey;
International University of Gorazde, Bosnia and Herzegovina

Abstract— Distributed ledger-based transaction management approaches and lineage data recording for dynamic management of DAG structures enable trusted scalability of analytical transactions in some manner. Submodular and disjoint cluster sets increase connectivity performance of analytical systems. Unified batch/interactive/adhoc querying can enable trusted E2E (end-to-end) analytics for massive systems. Trust factor and trust cost optimization reinforce the trustworthiness of the overall system in varying context. However, security and privacy architectures are still context dependent. The dynamic change in context have to be adapted to all system layers in real-time for a complete E2E trust mechanism and confidence in smart-ecosystems. In this study, we explored dynamic security and privacy aspects of E2E analytical transactions in varying context. Classical behavior modelling approaches such as cellular automata, chaotic systems, hierarchical block diagram modeling methods are cumbersome to adapt the dynamism. Persisting and ensuring the trust for varying contexts with an E2E trust mechanism enable to adapt the dynamism at massive scale. Initial arguments for data exchange over a hybrid-cloud node, instead of cell unit scenario in a simulated 5G environment with the trust mechanism, is evaluated. It is promising to meet zero latency requirement of MEC (Multi-access/Mobile Edge Computing) units. Transmitting data over a hybrid-cloud node rather than cell units can enable mobility of 5/6G ecosystem with the E2E trust mechanism.

Ziggurat: A Framework for Providing Scalability and Security in IoT Blockchains

Brett Glendenning, Riley Kiefer, Ashokkumar Patel
Department of Computer Science, Florida Polytechnic University, Lakeland, Florida, USA

Abstract— As the internet continues to grow, networked devices are expanding not only in number, but variety as well. The Internet of Things presents an opportunity to connect every-day devices in new and intelligent ways, whether those devices are as mundane as kitchen appliances, or as complex as Autonomous Vehicles. To ensure the security of information in these sometimes public-facing, geographically distributed networks, devices cannot trust one another blindly. A Blockchain protocol may be ideal in that it trades trust in individuals for trust in cryptography, but the performance of such protocols can suffer in densely populated networks. To provide a scalable solution to Blockchain implementations IoT networks, this paper proposes a framework that utilizes a hierarchically distributed Blockchain network with horizontal connections between select devices at each layer to balance traffic without creating a single point of failure.

A Preliminary Design of Privacy Ontology for Smart Toys

Otavio de P. Albuquerque, Farkhund Iqbal, Patrick C. K. Hung, Marcelo Fantinato
School of Arts, Sciences and Humanities, University of Sao Paulo, Sao Paulo - SP, Brazil;
College of Technological Innovation, Zayed University, Abu Dhabi, UAE;
Faculty of Business and IT, Ontario Tech University, Oshawa, Canada

Abstract—With the technologies progress, smart toys have become more important in the toys market. With different shapes and purposes, toys company adopt different requirements and Web Services to create the smart toy's features. Generally, each company has its requirements and implementation process, including semantic information and risk management guidelines. That is, there is no common knowledge base related to the smart toy domain, which on the organizations could share information and reuse a standard knowledge, mitigating interoperability issues. Our work aims to build a smart toy's privacy ontology, bringing general concepts and privacy-related, machine-readable, to the domain, in view of to provide to organizations and software agents a common knowledge base related to privacy on smart toy's context, to reuse to smart toys and features implementation.

CSCI-ISED:
EDUCATION - STEM, COMPUTER SCIENCE AND COMPUTER ENGINEERING

Redesigning an Undergraduate Course: Principles of Programming Languages

Xuguang Chen
Computer Science Department, Saint Martin's University, Lacey, Washington, USA

Abstract - This paper describes the effort of redesigning an undergraduate course: CSC370 Principles of Programming Languages. Like courses similar to CSC370, the book Concepts of Programming Languages and written by Robert Sebesta is as the textbook. The lectures sequentially cover the topics from its Chapters 1-10. Based on the students' feedback and their performance, this book is good, however, with several limitations. Hence, the lectures were redesigned in the Summer of 2021 to reflect the concerns. In the new design, the topics selected from the book will be reorganized into three categories, and the coursework will include two term projects for different purposes. It will be piloted in the Spring of 2022.

Cybersecurity Integration: Deploying Critical Infrastructure Security and Resilience Topics into the Undergraduate Curriculum

Stephanie Lunn, Monique Ross, Jason Liu
Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, Florida, USA

Abstract - Security is an essential consideration as technology continues to flourish and permeate society. To establish professionals trained to actively prevent and combat threats, educators must find ways to increase awareness, interest, and familiarity with cybersecurity topics. In this paper, we describe our approach to deploying Critical Infrastructure Security and Resilience modules into an existing undergraduate course. We applied a combination of the Communication Theory of Identity and Disciplinary Identity Theory, specifically focusing on Computing Identity, as the frameworks that guided our inquiry and interpretation of the findings. We also utilized pre- and post-module surveys (n= 51) and Wilcoxon signed-rank tests to assess the efficacy of the modules over three semesters, and to learn more about their impact on students understanding, attitudes, and interest in cybersecurity topics. Upon completion of the modules, students reported feeling more confident in ``undertaking and succeeding in'' finding ways to exploit vulnerabilities in existing software, and in implementing protocols to allow data to be sent securely over a network, among other findings. While intended to provide the benefits of embedding rigorous modules into the curriculum, we also discuss additional pedagogical approaches that can further develop the skills of the future cybersecurity workforce.

Evaluation Strategies at University for Students with Dyslexia: a pilot study supported by face emotion recognition

Maura Mengoni, Silvia Ceccacci, Noemi Del Bianco, Ilaria D'Angelo, Simone Aperecida Capellini, Catia Giaconi
Departiment of Industrial Engineering and Mathematical Sciences, Universita Politecnica delle Marche, Ancona, Italy;
Department of Education, Cultural Heritage and Tourism, Univerity of Macerata, Macerata, Italy

Abstract - The growing number of students with disabilities enrolled at university requires rethinking the educational and teaching proposals from an inclusive perspective. It implies careful planning of the most delicate phase of the teaching and learning process for all students: the final exam. It is full of expectations and anxieties. Special attention to constructing a welcoming environment becomes essential, especially for students with Specific Learning Disorders (SpLDs) or disabilities. Therefore, this contribution, starting from a pilot study conducted by the University of Macerata, analyzes the role of Emotional Feedback in the assessment procedures.

A Circular Framework from Branding to Recruitment in Perspective of Higher Education

Miaoqing Chen, Ali A. Al-Babawat, Abeer Alsadoon, Srujal Patel, Razwan Mohmed Salah, P.W.C. Prasad
Study Group Australia, Sydney, Australia;
University of Duhok, Duhok, Kurdistan, Region-Iraq;
Kent Institute Australia, Sydney, Australia;
Australian Institute of Higher Education, Sydney, Australia

Abstract— This paper aims to enhance alumni loyalty as time passes with the help of technologies and also to propose a model regarding how branding links to university recruitment through the mediating role of alumni loyalty. The proposed model added a new stage called situational assessment by separating the single student loyalty into current student loyalty and alumni loyalty to enhance alumni loyalty as time passes with the help of effective loyalty programs, leading to new student recruitment and university reputation via positive Word-Of-Mouth, and included positive academic experience and quality digital technologies to solve the problem of difficulty in measuring service quality as they play a significant effect on alumni loyalty as time passes in terms of new student recruitment and university reputation.

Modify Flipped Model of Co-regulation and Shared-regulation Impact in Higher Education, and Role of Facilitator on Student's Achievement

Aida Jafarian, Razwa Mohmed Salah, Abeer Alsadoon, Srujal Patel, Gustavo R. Alves, P.W.C. Prasad
Study Group Australia, Sydney, Australia;
University of Duhok, Cihan University-Duhok, Duhok, Kurdistan, Region-Iraq;
Kent Institute Australia, Sydney, Australia; Polytechnic of Porto, Porto, Portugal;
Australian Institute of Higher Education, Sydney, Australia

Abstract- Flipped learning approach is a well-organized learning model leading to efficient active learning, effective peer-to-peer collaboration, and student-teacher interaction. To date, strategic implementation of co-regulation and shared regulation is rare in a flipped model in higher education, but they are become necessary to apply them. This research is to propose and modify the current flipped learning model by adding some elements such as providing some co-regulation and shared regulation strategies to enhance the level of student's self-regulation skills to better student's academic achievements next to instructor support, and peer-topeer interactive way by creating a regulation questions to collaboratively assessment of the student's self-regulation resulted.

Designing an Extreme Based Teaching Model for the First Programming Course

Leon Pan
University of the Fraser Valley, Canada

Abstract - Due to the high demand for computer science related jobs, more and more students enroll computer science related courses, especially programming courses. Unfortunately, the first programming course is very challenging for many of them. There is usually much higher drop-off and failure rate in the first programming course. To deal with this problem, this paper proposes an extreme based teaching model which takes the advantage of merits of extreme programming methodology. The model integrates the lecture-based, project-oriented, and innovative learning. The teaching materials are organized into a number of units, each of which is promoted by a well designed accompanying project. Students working in groups can constantly communicate with and help each other and evaluate and improve the performance of each student and group. With the intensive involvement of the entire teaching and learning process, instructor is able to provide timely feedback and advice, continuously adjust and refine the future teaching unit(s). It is expected that students will learn more efficiently and effectively in this environment, which results in higher successful rate in the first programming course.

A Dev-Ops Practicum Course for Information Technology Majors

Adam Kaplan
Department of Computer Science, California State University Northridge, Northridge, California, USA

Abstract— Since the inception of the CSUN CIT major one decade ago, the IT industry has experienced two large paradigm shifts: from on-premise virtual servers to scalable cloud-hosted services, and from full virtual machine deployment to containerized application deployment. Many IT professionals are now dev-ops specialists. Rather than install and configure individual servers, dev-ops specialists program and maintain Infrastructure as Code (IaC). In this manner, application testing, deployment, and scaling can be orchestrated automatically to respond to changing network and service conditions. Such dev-ops skills are increasingly demanded of new Information Technology graduates. Thus, we have redesigned our experimental senior-level CIT project course to be structured as a dev-ops practicum. This course now comprises six hands-on projects in web application deployment, automated testing, source code repository management, cloud hosting, infrastructure as code, and application containerization. Student outcomes were extremely positive, owing to individual rather than team projects, live (virtual) lecture format as opposed to a flipped classroom, a broad set of project topics and practiced skills, and frequent one-on-one help sessions.

Authentic Assessment in a Beginning Cybersecurity Course: A Pilot Study

Glen Sagers
Computer Science and Information Systems, Southern Utah University, Cedar City, Utah, USA

Abstract—Authentic assessment means to examine student performance on real-world tasks. This study explores the process of developing and using authentic assessment in a basic cybersecurity course, specifically in place of written examinations. Although only a pilot study, initial results are promising. Each assessment was a learning experience, besides assessing a skill, and student perceptions of the assessment were positive. Some students also went beyond the bare minimums needed for the assessment, likely due to a sense of ownership of the IT resources. Using authentic assessments in place of exams also removed any need for exam proctoring.

Impact of COVID-19 and Social Measure on Behavioral Intention in the Context of Online Learning Software

Minje Cho, Hyeon Jo
Hospitality and Tourism Institute, Duy Tan University, Da Nang, Vietnam;
IT Laboratory, Realsecu, Busan, South Korea

Abstract—Since the COVID-19 outbreaks, countries around the world have implemented social measures to prevent spread. To inhibit infection, governments are conducting social distancing, wearing a mask, and vaccination. In particular, in the field of education, non-face-to-face lectures are being carried out, and the importance of online lectures has increased. The purpose of this study is to understand the behavioral intention of students in the context of online class software. Data were collected from 318 Vietnamese university students taking classes through online education software, and PLS analysis was performed. As a result, attitude has a significant effect on behavioral intention. Contrary to expectations, subjective norms do not affect attitude. Perceived behavioral control, regulatory environment, and affective risk perception are significantly related to attitude. The results of this study will provide meaningful guidelines for implementing effective education within COVID-19 in the future.

Learning Finance with Games: An Empirical Study

Sandy Ingram, Rania Islambouli, Miharisoa Hanitrarivony Andrianantenaina, Jean-Philippe Weisskopf, Philippe Masset, Nicole Baudat
HES-SO, HEIA-FR, Fribourg, Switzerland;
EPFL (The Ecole Polytechnique Federale de Lausanne), Lausanne, Switzerland;
HES-SO, EHL, Lausanne, Switzerland; HES-SO, Source School, Lausanne, Switzerland

Abstract—This paper presents an empirical case study on applying game-based learning in higher education, more precisely in an undergraduate finance course. The paper describes the experimental study context, protocol, and results. Using multivariate regression analysis, a significant game effect on student performance is observed for two of the three games explored. This suggests a positive influence of gamification on the learning experience and outcome.

Agile Transformation for Capstone Projects: Preparing Graduates for the Job Market

Maha Al-Yahya, Rana Alkadi, Hend Alrasheed
Department of Information Technology, King Saud University, Riyadh, Saudi Arabia

Abstract—Preparing graduates for the job market is a key objective of higher education. The Information Technology department at King Saud University has adopted a strategy of program alignment with industry to ensure that program outcomes are in line with the market needs and requirements. Graduates in the field of Information Technology should be equipped with software development skills needed by industry to drive business value and deliver high quality software products and services. To this end, the IT department undertook the decision to adopt an agile transformation strategy for the final year capstone project course converting it from a waterfall software development process model to an agile approach in response to the job market need. In this paper, we present the transformation strategy, the design of the course, and discuss opportunities and challenges. Reporting our transformation experience will provide insights and guidance to those who want to undergo a similar transformation.

Assessing Cognitive Load for Junior Software Engineers: A Mixed-Method Study

Mark Allison, Stephen Turner, Molly Kwansy
The University of Michigan - Flint, Michigan, USA

Abstract—The rising complexity of software systems combined with the proliferation of new tools and techniques has challenged Software Engineering educators to refine pedagogical techniques in preparing junior developers. Despite the growing body of literature devoted to computer science education in general, there exists only scant studies from a software engineering practice perspective. In this paper we examine this concern through the lens of perceived cognitive load. A mixed mode study was conducted within a senior level capstone course. Our results suggest that while significant cognitive load was experienced throughout an agile process (design, project management and implementation), participants found the latter most demanding. There were indicators that mitigating strategies mindful of working memory could be effective, and there exists a significant requisite to scaffold the skillsets applied in navigating documentation, and determining appropriateness for task requirements.

Reinforcement Learning - Based Control Systems for Networked Power Infrastructures

Namrata Saha, Shabnam Rezapour, M. Hadi Amini
Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, Florida, USA;
Enterprise and Logistics Engineering, Knight Foundation School of Computing and Information Sciences,
Florida International University, Miami, Florida, USA

Abstract—Coordination in today's large power systems is critical to satisfy demand fluctuations and reduce the required generation capacity. In this paper, we employ Reinforcement Learning (RL) techniques to design a control system to enhance coordination among facilities and minimize the power production cost. The controlling agent of the RL observes the the system's state (e.g., demand volumes), perform actions (e.g., production volumes), and get rewards or punishments as the effect of their actions. Therefore, agent gets trained in the environment, find an optimal policy towards the system's goal, and establish a smart energy system. The developed control mechanism is tested for different loads (residential, commercial and industrial) and compared against non-smart approaches that exist in the literature.

Evaluation of Web-based Digital Storytelling Tools for Early Childhood Education

Marianthi Manousi, Jim Prentzas
Department of Education Sciences in Early Childhood, Democritus University of Thrace, Alexandroupolis, Greece

Abstract—ICT tools play an increasingly important role in all educational settings. During the last decades, ICT tools have been incorporated in the curriculum of early childhood education. An ICT resource gaining interest in early childhood education involves digital storytelling. Storytelling always played an important role in early childhood. It is expected that digital storytelling will provide additional benefits. A missing aspect in related work is the use of alternative Web-based tools for the creation of digital stories in early childhood education. This aspect was dealt in a recent approach presented by the authors. In this paper, the alternative Web-based authoring tools are assessed using two alternative models. This assessment will assist teachers, researchers and developers. To the best of our knowledge, the specific discussion is missing from related work.

Undergraduate In-class Research Experience for Computer Architecture Students

Muhittin Yilmaz
Electrical Engineering and Computer Science Department, Texas A&M University-Kingsville, Kingsville, Texas, USA

Abstract—This study presents a hands-on research experience for undergraduate senior-level computer architecture course students. The students have investigated scientific research process, literature review approaches, technical writing as well as blind-review principles, and conducted hands-on research on three different computer systems, namely, a supercomputer, an office desktop, and an autonomous vehicle artificial intelligence computer systems, for a budget-constrained final computer configuration of an office desktop computer. The final student team outcomes, relevant feedback, and the corresponding surveys, evaluated by the project administrators, strongly imply the success of the project for an effective research component inclusion in an undergraduate course.

MarieSimR: The MARIE Computer Simulator Revision

Xuejun Liang
Department of Computer Science, California State University - Stanislaus, Turlock, California, USA

Abstract – MarieSimR, a revision of MarieSim [1], is developed to support the use of stack. To this end, the stack pointer stored in a reserved memory location and the stack-relative addressing mode are added. The MARIE subroutine call and return instructions are revised to use the stack for saving and getting the return address, respectively. A stack frame can be created for a subroutine to hold the return address, input arguments, output results, local variables and so on. Therefore, recursive subroutines can be supported. A new instruction for increasing or decreasing the value of the stack pointer is added to facilitate the push and pop operations. In addition, a new instruction for loading an immediate constant into the accumulator is also added to replace the MARIE clear instruction. Finally, a new assembler directive is added to support for defining a label to hold the address of another label symbolically. In this paper, the design and implementation of this revision is presented. Two programming applications are also discussed to illustrate how to utilize these new instructions.

A Real-Time Translation of Arabic Video Contents to Arabic Sign Language

Ashwag Maghraby, Safaa Qutub, Afraa Alandijani, Huda Alandijani, Linah Bakhsh, Nuha Alharbi
Department of Computer Science and Informatics, Umm Al-Qura University (UQU), Mecca, Saudi Arabia

Abstract—The shift to online education due to the COVID-19 pandemic has brought various challenges to different individuals. Among these affected individuals are deaf and hearing-impaired. They face difficulties in understanding online videos because they do not provide real-time translation into Sign languages. The main idea of this project is derived from this point. It proposes a PC-based application called "Al-Banan" that supports deaf and hearing-impaired in the Arabic region to understand media in general including learning video and live virtual classes by translating Basic Arabic Speech into Arabic Sign language using a three-dimensional avatar. The application was built done using Python and C# languages. The application was tested on many deafs through social media and proved to be effective if the word exists in the dictionary a three-dimensional cartoon avatar showing the corresponding Saudi Arabia Sign Language. Otherwise, the word is represented by fingerspelling. The results proved the understanding and acceptance of the idea by the deaf.

Securing On-Chip Communications: An On-The-Fly Encryption Architecture for SoCs

Amin Sarihi, Ahmad Patooghy, Mahdi Hasanzadeh, Mostafa Abdelrehim, Abdel-Hameed A. Badawy
New Mexico State University, Las Cruces, New Mexico, USA;
North Carolina A&T State University, Greensboro, North Carolina, USA;
Independent Researcher, Tehran, Iran; California State University, Bakersfield, California, USA

Abstract—This paper proposes an on-the-fly encryption system that consists of a lightweight encryption architecture and a keyexchange algorithm for secure on-chip communications. The hardware-optimized encryption architecture protects chip-wide communications against a malicious router by applying fast data encryption on packets of interest. The key-exchange algorithm is used to ensure that encryption keys are fully protected from the malicious router. The proposed security solution establishes a trade-off between the achieved security and the imposed overheads through adopting the number of encryption rounds. The two versions of the proposed encryption system, i.e., a besteffort and a full-protection, respectively, offer 99.5% and 100% data protection against the target security threats. While our RTL-level evaluation shows that the maximum area overhead is 19% relative to a baseline router, our network-level simulations show almost no performance overhead.

Designing Diversity, Service, and Active Learning for the Common Good in an Undergraduate Computer Science Course

Arshia Khan, Kathryn Milun
Department of Computer Science, Swenson College of Science and Engineering,
University of Minnesota Duluth (UMD), Duluth, Minnesota, USA;
College of Liberal Arts, University of Minnesota Duluth (UMD), Duluth, Minnesota, USA

Abstract - According to the literature, Computer Science courses are highly technical but are lacking in soft skills that enable students to develop not only the technical skills but also professional skills such as collaboration, communication within and across cultures, teamwork and knowledge of diverse cultures. In this paper we report on the design and execution of a prototype framework for using active learning to build soft skills and diversity awareness into a service learning component for a database management course. In this course students were given an opportunity to work with a regional Native American organization to create a database solution for a maple syrup operation that needed greater coordination for marketing, distributing and delivering their product across ancestral homelands disrupted by colonialism. Our aim was to test the prototype framework to see if it increased student learning over the traditional final exam and brought greater self-awareness, recognition and respect concerning cultural difference. Students reported that they learned more about the technical aspects of database creation from doing the service learning project than from preparing for the final exam; they also reported greater awareness of cultural diversity. This paper lays out the design of our service learning pedagogy, the test results, and the model we propose for other computer science technical courses.

Computer Science Theory from a Practical Point of View

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

Abstract—For the past several years, we at Baylor Computer Science, have been teaching a course entitled "Foundations of Computer Science." The objective of this course is to teach practical applications of mathematical machines, especially finite state machines and push-down automata. Undecidability and NPCompleteness are also covered. This course does not emphasize proofs and is not a substitute for a standard theory course. Instead, it is designed to give basic information about theoretical machines and their practical applications.

The Effect of Web-toon Content on Ego-identity and Ego-Resilience: Focused on Middle and High School Students

In-Hee Cho, So-Young Kim, Ji-Yun Oh, Sung-Min Cho, Sung-Hoon Park
Dankook University, Media Strategy Institute, Chungbuk National University, Chungju, Chungbuk, Korea;
Department of Computer Engineering, Chungbuk National University, Chungju, Chungbuk, Korea

Abstract - This study was focuses on middle and high school student enjoying web-toon. We examined the contents of the web-toon with 5 items (entertainment, work, quality, story, writer). And the effect of the content of Web-toon on ego - identity and ego - resilience. The survey was conducted from May 7, 2021 to May 21, 2021 for middle and high school students enjoying the web-toon of residents in Seoul. A total of 200 questionnaires were distributed, with the exception of seven questionnaires who responded unfairly. A total of 193 valid questionnaires were used as the final analysis. The results of this study were as follows. All five items(entertainment, work, quality, story, writer) of Web-toon content have positive effects on ego identity and ego - resilience. The following suggestions. To the general characteristics of entertaining and quality. We have diversified the factors of content quality by adding new items such as work, story, and artist. In addition, we confirmed that Web-toon is an important medium that mentally influences youth growth. Future research will focus on how web-toon influence adults and how to additivity of middle and high school students to web-toon through smart phones.

Usability Heuristics for Early Primary Children: A Case Study in Sri Lanka

S.M.U.P. Samarakoon, T. A. Weerasinghe, Hakim Usoof
University of Colombo School of Computing, Colombo, Sri Lanka;
Department of Statistics and Computer Science, University of Peradeniya, Peradeniya, Sri Lanka

Abstract— With the encroachment of technology, tabletmediated learning is widespread among kindergarten children. With this trend, most educational app developers target preschool children in their app market. Hence, developers are increasingly creating educational apps that target this age group. Similarly, usability plays a major role in pre-school children targeted applications due to lack of cognitive and physical development of this age. However, as per previous researchers, most of the current applications for pre-school children are not successful due to a lack of usability. Hence, preschool children face lot of usability issues when using these applications. Due to a lack of cognitive and physical development, children of early primary education find it difficult to use those applications. There are many usability frameworks, but most of them have been developed targeting adult users. So, there is no standard usability framework for early primary children. Hence, this research aims to propose a set of usability heuristics that can be used by the designers of tablet-mediated applications for early primary children.

The Use of Augmented Reality to Promote Learning in Science Laboratories

Francielli Freitas Moro, Maria Paula Correa Angeloni, Liane Margarida Rockenbach Tarouco, Joao Bosco da Mota Alves
Programa de Pos-Graduacao em Informatica na Educacao, Federal University of Rio Grande do Sul, Porto Alegre, Brazil;
Programa de Pos-Graduacao em Tecnologias da Informacao e Comunicacao,
Federal University of Santa Catarina, Ararangua, Brazil

Abstract—This paper presents a research on the use of Augmented Reality (AR) in Science teaching laboratories, aiming to verify features and functionalities that promote the interaction, participation and experience of students in line with learning theories, while supporting experiential learning. From the research, a prototype of a virtual laboratory involving the use of AR was developed, and a strategy for its application in learning activities about the planet Earth for elementary school students was built.

Design of a New Innovative and Systematic Education Program for Tech-oriented Startups and Scaleups Leaders

Dong H. Lee, Hwa Jeong, Seo Jinung An, Jong H. Lee
Center for Technology Startup Education Team, Research Institute of Convergence DGIST, Daegu, Korea

Abstract—we design a new TVA startup and scaleup program and implement the innovative technology venture education structure that can systematically train tech-based startups and scaleups leaders compared to existing startup business schools. It gives a good chance to acquire knowledge related to technology venture startups and obtain the ability to manage social enterprises and differentiated training program to provide customized education for each stage of corporate growth. In addition, through various technology innovation seminars and hybrid online and offline networking platforms, both startup and scaleup leaders who can promote two growth engines will be trained by sharing implementation strategies through individual business (or startups) presentations and exchanges, analyzing and consulting with experts. In the future, the TVA program keeps track of the 4th industrial revolution driving technology, develops related business models, and prepares an opportunity to challenge niche market.

Discovering Effective Learning Methods and Impact of Team-based Programming Projects in Graduate Level Courses

Dewan Tavir Ahmed
Department of Computer Science, University of North Carolina at Charlotte, North Carolina, USA

Abstract - Learning can be more efficient, effective and interesting if we can identify more about our students and know how they learn. Due to COVID-19, schools and colleges are offering online classes. It has a significant impact on students' success. Therefore, course modality and teaching pedagogy need to be taken into consideration for crafting and creating instructional experiences that make leaning appealing and effective. A number of innovative teaching methods such as active learning, hybrid learning, social learning and flipped classrooms have been proposed and tested. Practically, several methods together can be helpful for students. In this study, I conducted an experiment and identified effective learning methods for graduate level courses. According to this study, 94% students responded positively about this course design. The results also show that 83.5% students prefer face-to-face classes and 97% students find in-class problem solving effective to understand a concept better. Many courses incorporate team-based learning which is a proven approach. In this study, the benefits and limitations of team-based programming projects are identified as well as students' opinion in this regard. The results show that 85% students prefer team-based programming projects. Surprisingly 59.1% students mentioned all members do not contribute fairly evenly. This is a common problem in group works. So, small group size could be effective to overcome this problem.

Challenges of Incorporating a Mobile Device Forensics Summer Camp with Critical Thinking Amongst K-12 Students

Frances C. Dancer, April L. Tanner, Mary M. Thomas
Department of Electrical and Computer Engineering & Computer Science, Jackson State University, Jackson, MS, USA

Abstract— The researchers opted in a Mobile Device Forensics Summer Camp with middle school students from the surrounding areas of Jackson, MS, who debated on critical thinking, deductive reasoning, and problem-solving skills and defined the meaning of Mobile Device Forensics and what it entails in order to for the students to achieve proficient or advanced levels on the students state exams. The researchers had to overcome some barriers for the middle school students that subsequently had not occurred previously due to Covid-19, measure how effective this process progressed moving forward by the external evaluator, and discuss the challenges faced by the Leadership Team of the Mobile Device Forensics Group and how the researchers alleviated some of those issues.

Experiences with Hundreds of Similar and Customized Sites with DevOps

Christian Rodriguez, Geronimo Afonso, Lucas Di Cunzolo, Mauricio Foster, Leandro Di Tommaso, Fernando G. Tinetti
Mikroways, La Plata, Argentina;
Facultad de Informmatica, UNLP, CIC Prov. Bs. As., La Plata, Argentina

Abstract—In this paper, we present the experience and guidelines for deploying and managing hundreds of Moodle sites for providing services to teachers and students in hundreds of Argentinian schools. The selected e-learning platform is Moodle open-source learning platform, which is personalized and adapted due to the (huge) heterogeneity of the Buenos Aires province public school system at pre-primary (aka initial, in Argentina), primary, and secondary (pre-university/tertiary) education levels. The general approach is that of the well-known DevOps process, given that it lets a natural integration of source code development as well as the infrastructure needed for supporting the runtime daily processing. A Fraction of the e-learning system heterogeneity at different school levels is handled via source code. Heterogenous runtime daily operation requires automated infrastructure tasks management, such as workload balance, metrics collection and storage, performance visualization and alerting, backups, etc.

A Strategy to Assess Computational Thinking

Antonio Gonzalez-Torres, Lilliana Sancho-Chavarria, Magaly Zuniga-Cespedes, Jorge Monge-Fallas, Jose Navas-Su
Computer Engineering Department, Costa Rica Institute of Technology, Cartago, Costa Rica;
Computing School, Costa Rica Institute of Technology, Cartago, Costa Rica;
Omar Dengo Foundation, San Jose, Costa Rica;
Mathematics School, Costa Rica Institute of Technology, Cartago, Costa Rica

Abstract—The development of Computational Thinking is essential to improve the social conditions and employability of people. This type of thinking encourages the use of abstraction, planning, the decomposition of problems in the parts that constitute them, the learning of skills for specification, analysis and problem-solving. In addition, it contributes to the formation of critical thinking, creativity, and cooperation among students. However, there are no robust and scalable methods to study the enormous codebases that are produced by the solution of exercises, exams, and projects to recognize the predominant learning patterns and determine the fulfillment of objectives, the level of competencies achieved, and the goals accomplished. Consequently, it is necessary to carry out an exhaustive analysis of skills, knowledge, results, and levels of the performance of students to identify points of improvement to accomplish the learning objectives and amend the curricula and exercises. Therefore, this paper outlines the joint research under development by the Costa Rica Institute of Technology (TEC) and the PRONIE MEP-FOD program. It aims to automatically analyze the source code created by the students to assess the Computational Thinking skills they developed and compare the results with the learning objectives set in the curricula.

A Remote Instructional Approach with Interactive and Collaborative Learning to Teach an Introductory Programming Course during COVID-19 Pandemic

Mahmudur Rahman, Roshan Paudel
Morgan State University, Baltimore, Maryland, USA

Abstract—This work presents our experience and approach of teaching an introductory CS programming course remotely in this pandemic era without missing out on the benefits of hands-on learning generally used in-person learning environment. Synchronous virtual learning occurs when students join an audio/video enabled meeting space at the same time through several cloud based services, such as Zoom based conference tool, interactive cloud based coding environment in repl.it and GoSoapBox platform for online in-classroom engagement, which were well integrated in the Canvas based Learning Management System (LMS). To keep the structure of the session much like an in person learning experience, the synchronous session included whole group instruction in Zoom led by the instructor and small group (breakout room) based lab work in Repl.it amongst the learners. Both interactive and collaborative learning are infused in pedagogy effectively so that students can learn using interactive platforms, tools, technologies, systems, and services as available to them and collaborate within and among groups. To evaluate the impact of this infusion, a pre- and post-survey were conducted on student cohort (4 sections taught by 3 different instructors) in the Fall'2020 semester. In addition, final project scores and final grades for Fall'2020 semester and enrollment number and final grade distributions from Fall'2017 to Fall'2020 were also available for analysis. The initial evaluation of the survey results and student's performances based on quality point scores show evidence to conclude that the proposed pedagogical approach increased student motivation and engagement and facilitated learning to entry-level computer science students.

Undergraduate Research Experiences: A Case Study from the Computer Science Department, University Of Guyana

Penelope DeFreitas
Department of Computer Science, University of Guyana, Georgetown, Guyana

Abstract— The world is facing a wide range of serious problems that require urgent attention. Inherent to understanding and solving these problems are research and development activities. Research and development skills are primarily taught at institutions of higher education and have translated into numerous benefits for students, such as enhanced cognitive and personal skills. This paper presented a model for running undergraduate research courses at the Computer Science Department, University of Guyana (UG). Key components of the model include preparatory assessments, two-semester offering of research courses (formal proposal and thesis), human resources, mechanisms for guidance and support, assessments and grading rubrics. The model was first implemented in 2017/2018, and was monitored and evaluated over a two-year period using a sample of 23 undergraduate students who were pursuing computer related BSc degree programs. Students provided feedback via a survey that was primarily focused on the research process, supervision, course structure, weekly research sessions, and general strengths and weaknesses of the undergraduate research model. The participants reported positive feedback such as having a good grasp of the research process, high levels of confidence in conducting independent research and preparedness for postgraduate studies. As such, this undergraduate research model may have wider applicability for the UG community and other institutions of higher learning.

The Power of Mentoring Programs in Retaining Women and Black, Indigenous, Students of Color in Undergraduate Computing Majors

Kristina Kramarczuk, Jandelyn Plane, Kathryn Atchison, Maya Narayanasamy
College of Education, University of Maryland, College Park, Maryland, USA;
Department of Computer Science, University of Maryland, College Park, Maryland, USA

Abstract—This poster outlines the effectiveness of an undergraduate mentoring program offered through the Department of Computer Science at the University of Maryland in retaining undergraduate teachers/mentors (which we call Ambassadors) in science, technology, engineering, mathematics, and computer science (STEM-CS) majors. Each Ambassador is responsible for providing support for current university students or for K-12 computing outreach programs. Enrollment and survey data show that the Ambassador program is successful in retaining undergraduate computing majors in STEM-CS. Ambassadors indicated that they appreciated the community aspect of the Ambassador program and they felt empowered to solve diversity problems in STEM-CS.

The Mental Challenges of Emergency Remote Learning: UAE Engineering Students Case Study

Sara Suleymanova, Amjad Gawanmeh, Wathiq Mansoor
College of Engineering and IT, University of Dubai, Dubai, UAE

Abstract—Student mental health in higher education has been an increasing concern. The COVID-19 pandemic situation has brought this vulnerable population into renewed focus. Governments had to close several sections, including educational institutes and universities, around the world suddenly in March 2020. Hence, emergency remote learning was adopted as alternative and as an immediate response to the ongoing situation using whatever available online tools. As a result, both students and instructors were forced to adapt to this new situation. While there were several studies that addressed several issues related to preparation, contents, course delivery, readiness, etc., there are few ones that were intended to address the mental challenges resulted from the shift to emergency remote learning. The main motivation behind this study is to understand the mental challenges and effects of the sudden transformation into emergency remote learning considering engineering students as case study and improve the delivery and experience of learning for both the instructors and the students.

Supporting Undergraduates for Careers in Computing and Engineering with Scholarships and Supervision

Dulal C. Kar, Scott A. King, Dugan Um
Texas A&M University-Corpus Christi, Texas, USA

Abstract - The main goal of NSF's S-STEM program is to support low-income, talented students to be successful in education and career in STEM fields. In this work, we present our experience on an NSF S-STEM grant - Supporting Undergraduates for Career in Computing and Engineering with Scholarships and Supervision (SUCCESS) that recruited and supported 39 talented and financially needy undergraduate students with potential to succeed in academic pursuits. Each student was mentored and guided for curricular and co-curricular activities for engineering and computer science majors and relieved from financial burden of paying tuition and other expenses with an amount of $7,000 per year. There were 17 community college transfer students and 22 high school seniors recruited altogether. Among them were 11 Hispanic students and 13 female students. Our findings and results on the grant are important to prospective S-STEM PIs (principal investigators) who plan to recruit female, Hispanic, and/or community college transfer participants. Despite the support of the scholarship, seven students could not continue in the program since they failed to maintain the minimum GPA requirement as they were working extra hours to pay for living expenses. All dropouts were from the recruited high school seniors, and none were from the community college transfers. In the Hispanic student group, it is found that their overall GPA went down near the end of the S-STEM program while other groups' overall GPAs remained unaffected or improved over the course of the program. The female students were found to maintain a higher GPA than that of the male students. The overall GPA of the students recruited from high schools went down, and all the attritions were from the group of the recruits from high school. In various support services and activities, midterm mentoring was found to be helpful for students at risk to continue in the program and undergraduate research participation was found to be effective in seeking employment in leading industry by some of the students.

Digital Measurement of Spatial Ability using a Virtual Reality Environment

Jhasmani Tito, Tania Basso, Regina Moraes
University of Campinas, Limeira, Brazil; University of Coimbra, Coimbra, Portugal

Abstract—Many activities, such as reading a map and selforienting in the right direction, are linked with people's spatial ability. This skill allows an understanding of the space-object relationship and its mental transformations. Digital tools such as Virtual Reality allow us to simulate real-world conditions to carry out interactive experiments to training and measure people's spatial skills, particularly mental rotation and spatial orientation. This work defines a method to collect information about the level of spatial ability, and how to train it, using a virtual reality environment. The method was submitted to a pilot test to preliminary validation. Thus, by applying the method, the authors aim to expand the scope of traditional test from 2D paper-and-pencil to 3D virtual environment, optimizing the data collection process.

Posters about Women Pioneers in Computer Science to Break the Gender Gap: A Preliminary Method

Lina Garcia-Cabrera, Irene Pedrosa-Garcia, Jose-Maria Serrano
Department of Computer Science, EPS, University of Jaen, Jaen, Spain;
Auringis High School, Junta de Andalucia, Jaen, Spain

Abstract— With the aim of integrating gender and culture issues into Computer Science and Engineering curriculum, this work proposes the design of posters about Women pioneers in Computing Science to break the gender gap. This contribution justifies the use of illustrations due to their great pedagogical potential. In addition, this paper describes the selection process of the Women pioneers in Computing Science and the criteria adopted for the graphic design of the posters, considering conceptual and communicative aspects as well as graphic, stylistic and aesthetic aspects.

Industry Connect Initiative: Partnering for Student Success

Bonnie Achee, Ghassan Alkadi, Matthew McNulty, Sandy Summers
Department of Computer Science, Southeastern Louisiana University, Hammond, Louisiana, USA;
Office of Career Services, Southeastern Louisiana University, Hammond, Louisiana, USA

Abstract— The Industry Connect Initiative of the Computer Science Department at Southeastern is a four-pronged approach to connect students to real world ready skills and relevant topics. This approach includes an industry advisory board, the distinguished lecture series, an internship program and curriculum opportunities. https://www.southeastern.edu/acad_research/depts/comp_sci/industry_connect/index.html These four pieces along with a partnership with the university's Career Services Department and Workforce Talent Initiative provide the resources necessary for graduates to be highly sought after for employment. This poster paper presents an overview of the Industry Connect Initiative

Application of Consumptive Metrics to Measure Internship Alignment

Mathew Nicho, Emad Batineh, Tarannum Parkar
College of Technology Innovation, Zayed University, Dubai, United Arab Emirates, UAE

Abstract—Internship aims at transitioning students from the academic environment (academic learning at the university) to a professional work environment (industry practice). In this respect, our paper aims at objectively evaluating the alignment of learning with practice based on the internship program conducted in term 1 2020 (pre-Covid) for our undergraduate students at the College of Technology Innovation studying bachelor's program in Computer Science and Information Systems. In order to measure the alignment, from a theoretical perspective, we adopted the framework of Kirkpatrick that provides a set of 'consumptive metrics' for evaluating the learning resources consumed in education and training using two constructs namely 'reaction' (how the learners feel, including their personal reactions to the internship training) and 'learning' (measure the knowledge, skills, or attitudes acquired as a direct result of the training including mapping to their courses). Using 36 internship student reports collected over a single semester (where students spent 8 weeks onsite at various organizations in the United Arab Emirates) as a sample for this study, we measured internship results in terms of the learning resources consumed during the internship experience using consumptive metrics to observe its alignment in practice. The results of the study provide academics to reinforce areas of strength and uplift areas of concern to align 'learning' with 'practice'.

The Transformation Process from In-campus Classes into Online Classes due to COVID-19 Pandemic Situation - The Case of Higher Education Institutions in Kosovo

Erza Baftiu, Krenare Pireva Nuci
Faculty of Computer Science and Engineering, UBT - Higher Education Institution, Prishtina, Kosovo

Abstract—The COVID-19 pandemic has caused changes in terms of traditional teaching globally. In Kosovo context, the Universities have found the transition from teaching in class to online classes quite challenging. This study investigates the transformation process from in-campus classes to online classes from the technical perspective within five Higher Education Institutions (HEI) in Kosovo. The data was collected using the qualitative methods and its analysis followed the 3C Litchman approach. The results show that each of the Universities followed a different approach by using either their limited premises infrastructure or using additional cloud infrastructure.

Statistical Study and Intelligent Classroom Proposal Adapted to Students with Speech and Hearing Disabilities

Asmaa El Mahmoudi, Habiba Chaoui, Nour El Houda Chaoui
Advanced Systems Engineering Laboratory, National School of Applied Sciences, Ibn Tofail University, Kenitra, Morocco;
National School of Applied Sciences, Sidi Mohamed Ben Adbedllah University, Fes, Morocco

Abstract - Modern teaching methods increasingly use new technologies. Students expect new and inventive ways of studying. For that, moving towards the design and development of smarts universities and classrooms is one of the most promising techniques, taking into account different approaches for students with disabilities (SWD). In this paper, we present a statistical study of the situation of students with special needs in the Moroccan university. Then, we present the state of the art of the various projects carried out by other countries. After that, we propose a dedicated smart classroom configuration for Moroccans SWDs, especially for deaf and dumb Moroccan students wishing to enroll in higher engineering schools.

Teaching Undergraduate Software Engineering: Xamarin Mobile App Development During the Covid-19 Pandemic

M. Javed, M. Estep
Department of Computing and Technology, Cameron University, Lawton, Oklahoma, USA

Abstract—The CS 4204 Software Engineering instructor was tasked with finding distance alternatives in teaching undergraduate face-to-face lab sessions in the fall 2020 semester. This was due to university policies established during the Covid-19 pandemic. This study describes teaching the course lab sessions utilizing Xamarin mobile app development on Mac and PC computers, with deployment to iOS and Android emulators and devices. Xamarin was chosen for flexibility as it could be used for both face-to-face and distance student learning. The teaching process is demonstrated in using stock and modified Xamarin tutorials which lead to a culminating mobile app development project. Student and instructor perceptions are discussed surrounding issues of facilitating software development instruction during the pandemic.

Impact of Covid-19 Pandemic on Higher Education

George Feuerlicht, Marek Beranek, Vladimir Kovar
Unicorn University, Prague, Czech Republic

Abstract - The sudden lockdowns in response to the COVID-19 pandemic forced educational institutions to rapidly transfer their teaching activities online. Border closures are having a major impact on universities around the world, in particular those with large cohorts of international students that are not able to attend their courses. In this paper we contrast the impact of the COVID-19 pandemic on the Higher Education institutions in Australia and the European Union and discuss the challenges and opportunities that this crisis presents. While most universities in Australia and the EU have implemented some form of online learning prior to the onset of the pandemic, it is now clear that the pandemic acted as a catalyst for rapid acceleration of digital transformation of the Higher Education sector. While many universities will embrace this transformation as an opportunity, others will resist major changes to their existing business models, and this inertia may be detrimental to their long-term success. Our analysis indicates that a successful transition to the post-COVID educational environment will be mainly determined by the technological readiness for online learning, the university business model and entrepreneurship of university management. We conclude that entrepreneurial leadership will play a key role in this transition.

A Survey of Innovation in Computer Science Education with Undergraduate Emphasis

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

Abstract - The focus of Computer Science Education(CSE) research has been innovation and improving the learning experience. Knowledge has grown exponentially in all disciplines within computer science. Over the last decade, computer science education has been evolving from abstract subject matter to increased innovation. These innovations have transformed the CSE from lecture to visualization, animation, simulation games, and hands-on experiential learning. Studies show improved understanding using automated and visual learning tools like simulations, virtualization, visualization, and video games. Recent research shows enhanced learning tools like visual automation, simulation, and video games are more beneficial than detrimental. This review will present some of the more innovative ways to give CSE. Even though the purpose is to focus on undergraduate CSE, the earlier studies originated for grade school and high school.

A Blended Learning Approach Supporting the Improvement of Primary School Students' Performance

Stavros Tsetsos, Jim Prentzas
Department of Education Sciences in Early Childhood, Democritus University of Thrace, Alexandroupolis, Greece

Abstract—The purpose of this study was to use a blended learning approach in the fifth grade of a primary school and investigate the impact on the students' performance in five subjects (i.e., Language, Mathematics, Science, Geography and History). Open educational resources were used. The duration of the research was approximately four months. The performance of students was better compared to the performance of students in another class of the specific school that learned with traditional instruction. The results of the quantitative data analysis showed that the post-test students' performance in Language, Mathematics and Science was statistically significant. On the contrary, students' performance in Geography and History seem to be statistically not significant. To the best of the authors' knowledge, there is no other reported blended learning approach in primary education involving five subjects with the specific duration and with results involving students' performance.

Student Hackathons for Low Cost Digital Solutions

German Terrazas, Gregory Hawkridge, Duncan McFarlane, Mike McNally
Institute for Manufacturing, University of Cambridge, UK;
Institute for Advanced Manufacturing, University of Nottingham, UK

Abstract—Academic research projects often face challenges when trying to deliver innovation in a fast and creative way. In particular, projects of multidisciplinary nature that look at cross domain aspects applied in industry may take longer to expect. This paper addresses such an issue in the area of low cost digital solution development and proposes an approach using student hackathons as a mean to enable rapid prototypes of low cost digital solutions for manufacturing SMEs. It highlights how these fast-paced activities could be seen as additional options to help students foster technical and soft skills ahead of working in industrial environments.

Evaluation of Distance Learning and Students' Achievement in Saudi Higher Education Institutions (HEIs)

Fahd H. Almutari, Sung-Chul Hong
Department of Computer and Information Sciences, Towson University, Towson, Maryland, USA

Abstract -This paper explores the distance learning process in Saudi Higher Education (HEIs) compare with in person learning and the COVID-19 pandemic effects on Saudi HEIs. The online survey examined the relationship between the distance learning process and students' achievement. The participants were from different Saudi HEIs faculty members, and based on their viewpoints, there is no important difference between students' achievement level in distance learning sessions and in person sessions. The COVID-19 pandemic makes the technology an essential option and could help for continuing higher education. The faculty members should be prepared to use technology and online platforms in lectures and communication with students and other staff. The authors found that there is no difference in faculty members' experiences when they used online platforms during distance learning process.

Effects of Social Distancing Intention, Affective Risk Perception, and Cabin Fever Syndrome on Perceived Value of E-learning

Hyeon Jo
IT Laboratory, Realsecu, Busan, South Korea

Abstract—Universities have switched from traditional face-toface lectures to e-learning to prevent the spread of COVID-19. Elearning provides the health value of prevention of infection along with the purpose of education. In this context, this study aims to identify factors affecting the perceived value of e-learning. It assumes that affective risk perception, cabin fever syndrome, and social distancing intention have a significant effect on the perceived value of e-learning. A comparison between groups between male and female was also performed. A questionnaire was conducted on university students and samples were analyzed using the PLS technique. The findings showed that social distancing intention determines perceived value. The analysis results revealed that affective risk perception has a significant positive effect on perceived value and social distancing intention. In addition, the cabin fever syndrome harms the value. There was a significant difference between male and female groups in the influence of social distancing intention on perceived value. Moreover, there was a significant difference between men and women in the path coefficient of affective risk perception on perceived value. Academic contributions and practical implications are discussed.

Reason for Ethics in Technology

Gary L. Wallace
Southern Utah University, Cedar City, Utah, USA;
(Retired) Commander and Director of Information Technology and Security, United States Navy, USA

Abstract - Recognizing the importance of ethics in industry today, the relevant strategy literature offers several anecdotal and prescriptive narratives to strengthen certain ethical concepts. The National Business Ethics Survey shows that employees have lost faith in management.s ethics. The industry has an ethical crisis of massive proportion and because of this; companies such as Enron have toppled taking employees. livelihood and life savings by crumbling down with the corporation. There appears to be no check and balance. The corporate culture has become engulfed with greed. Most of the current literature provides neither clear information nor an organizing framework to successfully tackle ethics in business. Currently, ethical rules and standards appear to be a .management only. club. To change this aspect in the industry will require the education of the employees. As employees are promoted or required to make ethical decisions, the basis of the decisions will come from ethics; therefore, ethical training is required. Companies seem to avoid training in ethics due to labor possibly having too much education will undoubtedly cause more overhead for organization in the way of questioning specific internal actions.

Hyperparameter Optimization and Comparison of Student Performance Prediction Algorithms

Ibrahim A. Alharbi, Ali Jaber Almalki, Cliff C. Zou
Department of Computer Science, University of Central Florida, Orlando, Florida, USA

Abstract—Educational Data Mining is the process of extracting information from datasets with educational relevance. This process of information extraction can be extremely insightful and useful for policy decision making but also small scale interventions. In our research, we optimize model configurations and hyperparameters of student performance prediction pipelines. Our models target to predict student performance based on data gathered in Portuguese schools pertaining to the subjects of Mathematics and Portuguese. The target variable is either cast in two or five bins and we train separate models for each of the tasks. We search model configurations amongst three different feature selection algorithms and 7 different classifiers implemented in scikit-learn. Our search spans 2000 iterations overall different experiment setup permutations. We successfully develop novel model configurations that perform exceptionally well. Furthermore, we set an important precedent showing the utility of hyperparameter selection and model search for Educational Data Mining.

Discussion Visualization based On Analysis of Comments to Web News

Hiroki Wakabayashi, Hiroki Nakayama, Ryo Onuma, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura
Department of Computer Science and Mathematics, Fukushima University, Kanayagawa, Fukushima, Japan;
Global Education Center, Waseda University, Shinjukuku, Tokyo, Japan;
Department of Computer Science, College of Liberal Arts, Tsuda University, Kodaira, Tokyo, Japan;
Division of Natural Science, Tokyo Gakugei University, Nukuikita, Koganei, Tokyo, Japan

Abstract—Opportunities for conducting discussions and expressing opinions on the Web environment have increased. In such discussions, it is desirable to fully understand the processes and main points of discussions. Especially, it is important to understand both the main points of discussion and trends of opinions together rather than superficial results. However, such understanding is often difficult since the number of utterances increases as time passes and discussions become complicated. In this research, we were aiming at developing methods for visualizing discussions in comments to Web news so that novices can experience the understanding both the main points of discussion and trends of opinions. This paper mainly overviews the prototype of discussion visualization system.

Design of Network Forensics Labs for Teaching-oriented Institutions

Kewei Sha, Wei Wei, Arpit Jain
Department of Computing Sciences, University of Houston-Clear Lake, Houston, Texas, USA

Abstract—Network-related cyber crimes including Phishing attacks, DDoS attacks, Identity Theft, etc. increase significantly with the extension of networked devices and systems. Cyber crime investigation conducted by network forensic professionals is critical not only to discover the source of security attacks but also to prevent future crimes. Although colleges and universities start involving forensics in education programs and forensics professionals training based on well-designed and effective curricula, there is still a significant gap between the supply of qualified network forensics professionals and what the market demands. As a highly technical subject, network forensics requires practitioners to obtain necessary knowledge and skills through both theoretical learning hands-on labs. However, survey of existing curricula shows the lack of hands-on network forensics labs. Therefore, we aim to develop a suite of initial hands-on network forensics labs that can be easily integrated into a Network Forensics course. The design focuses on addressing two major issues, building an isolated lab environment from the existing campus network and creating labs that provide realistic and practical experiences. In this paper, we first discuss the challenges of developing network forensics labs, then we describe our approaches to overcome those challenges. In addition, we also present the design of three typical network forensics labs. We believe that our designed labs, as well as the lessons learned, can help other institutions to develop effective network forensics courses.

Lessons and Experiences From Teaching Computing Science Capstone Project Courses

Kewei Sha
Department of Computing Sciences, University of Houston-Clear Lake, Houston, Texas, USA

Abstract—Capstone projects are bridges that connect institutional education and real-world industry settings. In the Capstone project course, a group of students work as a team on a real-world project that is usually provided by local industry. Instructors of Capstone project course work together with industry mentors to guide the Capstone teams to complete a team-long project beginning from requirement analysis. The success of the project depends on many factors. It is a challenge for an instructor to help the Capstone students to reach all the goals of the project. In this paper, we share lessons and experiences learned from the practices of teaching the Capstone project for many years.

Benchmarking the Software Engineering Undergraduate Program Curriculum at Jordan University of Science and Technology with the IEEE Software Engineering Body of Knowledge (SWEBOK V3.0)

Moh'd A. Radaideh
Department of Software Engineering, Jordan University of Science and Technology, Jordan

Abstract - This paper evaluates the compliance of the Software Engineering Undergraduate Program (SWE-Curriculum) at Jordan University of Science and Technology (JUST) with the last five of the fifteen Software Engineering Knowledge Areas (SWE-KAs) of the SWEBOK-V3.0 of the IEEE Computer Society. This research is the first to measure the coverage of the SWE-KAs in any SWE-Curriculum.

Shifting the Paradigms from Teaching Project Management to Teaching Software Project Management at the Jordan University of Science and Technology According to the IEEE Software Engineering Management Knowledge Area

Moh'd A. Radaideh
Department of Software Engineering, Jordan University of Science and Technology, Jordan

Abstract - Software Engineers usually find themselves in Project/Administrative/Operational Management positions. Therefore, it is very important to include a senior Software Project Management Course in any Software Engineering Curriculum. Software Engineers mostly experience Planning Challenges when it comes to Budgeting, scheduling, identifying and assessing potential Risks, and communicating with their Customers and Team-Members. Transforming the traditional Project Management Education to a Software Project Management Education is crucial to equipping Software Engineers with adequate Practical Experience in the Software Domain. This paper describes a senior Software Project Management Course that is based on the Project Management Body of Knowledge (PMBOK Guide), and the IEEE Software Engineering Body of Knowledge (SWEBOK-V3.0).

Estimation of Problematic Portions in Presentation Slides Based on Analysis of Audience's Unconscious Reactions and Slide Data

Ko Saito, Ryo Onuma, Hiroki Nakayama, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura
Department of Computer Science and Mathematics, Fukushima University, Fukushima, Japan;
Department of Computer Science, College of Liberal Arts, Tsuda University, Tokyo, Japan;
Global Education Center, Waseda University, Tokyo, Japan;
Faculty of Education, Tokyo Gakugei University, Tokyo, Japan

Abstract— Opportunities to make presentation using slides have increased to explain the results and progress of work. Therefore, there is an increasing need to develop the ability of students to improve presentation slides. To make students experience in refining slide by themselves, it is important to suggest the problematic portions in slides to them. In this research, we have developed a mechanism for automatically estimating problematic portions in a presentation based on the analysis of the compositional features of slides and gaze movements of audience. We have also been working on the development of a system that presents the results of the estimation in a understandable form so that students can use them in their refinement of slides. This paper mainly describes the features of a system that estimates the problematic portions in slides.

CSCI-ISPC:
SIGNAL & IMAGE PROCESSING,
COMPUTER VISION & PATTERN RECOGNITION

Neighborhood Base Matched Morphological Filters: Cross-fertilization with Linear Lowpass Filtering

Mahdi Khosravy, Neeraj Gupta, Olaf Witkowski, Antoine Pasquali
Computer Science and Engineering Department, Oakland University, Rochester, Michigan, USA;
Cross Labs, Cross-Compass Ltd., Tokyo, Japan

Abstract—The manuscript introduces Neighborhood Base Matched Morphological Filters (NBM-MF) by fusion of linear lowpass filtering process into the mathematical morphology structure of the operators. This internal cross-fertilization is gained via deployment a dynamic structuring element which adaptively matches itself to the neighborhood base of the signal/image. The neighborhood base is indeed the base-line/base-surface of the signal/image approximated by linear lowpass filtering called hereafter 'base'. This cross-fertilization enables NBMMF for modifying geometrical features of the signal/image more adaptively to the local geometry, and thereof having stronger filtering efficiency and less side effects of disturbing the original structure of the signal/image. The morphological smoothing is deployed for the efficiency evaluation of NBM-MF compared to the classical MF. Using three different numerical evaluation criteria, the morphological smoothing of signals with different noise and different base structure approves the higher efficiency of NBM-MF with respect to classical one.

Application Note: uPolar - An Interactive 2D Visualization Tool for Microscopic Time-series Images

Mehran Ghafari, Daniel Milman, Li Yang, Weiwei Dang, Hong Qin
University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA;
Baylor College of Medicine, Houston, Texas, USA

Abstract—To address the challenge to visualize time-lapse microscopy results, we have developed a circular plotting software tool, µPolar, to describe the trends and patterns of the cell movements and cell division events in a time-series. µPolar is interactive and easy to use. We demonstrate the utility of µPolar by visualizing the oscillating patterns of events of dividing yeast cells and the changing cell shapes of migrating mouse fibroblasts. µPolar potentially could be applied to other types of time-series of microscopic images. This R package µPolar is available through GitHub.

Gain and DC Offset Mismatch Correction in TI-ADC System Using Adaptive Filter

Muktiben Mehta, Sung-won Park
Department of Electrical Engineering and Computer Science, Texas A&M University-Kingsville, Kingsville, Texas, USA

Abstract− For digital processing of signals, Analog to Digital Converters (ADCs) are used to convert analog signals to digital signals. The sampling frequency of the ADC should be at least two times the highest frequency of an analog signal. To increase the sampling rate for super high frequency signals, time interleaved ADC (TI-ADC) systems are used. In a TI-ADC system, M ADCs are time interleaved to increase the sampling frequency by factor of M. Ideally all ADCs should have the same gain and dc offsets. However, practical ADCs that are interleaved have different gains and dc offsets which result in mismatch. This mismatch in turn gives spurious peaks in the frequency spectrum. In this paper, least mean square (LMS) and recursive least square (RLS) adaptive filters are used to correct the mismatch. The simulation results show that at least 20 dB attenuation in the spurious peaks after correction using adaptive filters.

A Review Paper on Facial Recognition Techniques in E-business

Tiyani Christopher Hlongwane, Topside E. Mathonsi, Deon Du Plessis, Tonderai Muchenje
Department of Information Technology, Tshwane University of Technology, Pretoria, South Africa

Abstract— Facial recognition is a biological biometric feature that allows a person to be identified from a digital image. The face is the most recognizable aspect of human anatomy and acts like a human being's first distinguishing feature. There are different techniques that can be used for the classification of data, two widely used techniques for data classification and dimension reduction are Principle Components Analysis (PCA) and Linear Discriminant Analysis (LDA). Facial recognition techniques have been comprehensively studied and applied in e-business. To reduce the False Rejection Rate (FRR) and False Acceptance Rate (FAR) during the recognition process, this review looks at the methods and the parameters that affect the face recognition. Furthermore, we outline the strengths and challenges of these techniques. This comprehensive study serves as a starting point and a guide for everyone interested in exploring in facial recognition techniques research area. Finally, the paper presents the conclusion and future work.

On the Selection of Hyperparameters in Convolutional Neural Networks

Donglin Wang, Qiang Wu
Department of Mathematical Sciences, Middle Tennessee State University, Murfreesboro, Tennessee, USA

Abstract—Convolutional neural networks (CNNs) have been showing great success in a variety learning tasks, especially object detection and image recognition. In the applications of CNN models, many hyperparameters, such as kernel size, channel size, learning rate, padding, and stride step, should be tuned and the success of CNN models highly depends on the correct selection of these parameters. However, it is very time-consuming and challenging to tune hyperparameters. In most cases one has to preselect some values for the hyperparameters based on experiences. In this paper we explore the impact of kernel size and channel size on the accuracy of feature extraction and imaging classification by CNN models and provide a practical guide to set up these hyperparameters. Particularly for kernel size we found the choice could be quite flexible for an easy task while the best choice should be in a range for a difficult task, and the range shrinks as the feature to background ratio becomes smaller.

Automatic Classification of Coronary Stenosis using Feature Selection and Boltzmann-Univariate Marginal Distribution Algorithm

Miguel A. Gil-Rios, Igor Guryev, Juan M. Lopez-Hernandez, Ivan Cruz-Aceves
Sciences and Engineering Division, Universidad de Guanajuato, Salamanca, Mexico;
CONACYT - CIMAT, Centro de Investigacion en Matematicas, A.C., Guanajuato, Mexico

Abstract—In this paper, a novel method for automatic classification of coronary stenosis in X-ray angiograms is presented. The method consists on the steps of automatic feature selection driven by a Boltzmann-Univariate Marginal Distribution Algorithm and a Support Vector Machine for classification. In the feature selection step, a set of 49 computational features was determined for working with a database of 2788 coronary stenosis images. According to the experimental results, the proposed method achieves a classification rate of 0.86 and 0.75 in terms of the Accuracy and Jaccard metrics, which is highly suitable to be considered in computer-aided diagnosis.

Efficient 2D Keypoint-based Hand Pose Estimation

Shan-Chien Hsiao, Ching-Te Chiu
Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan

Abstract—In this paper, we aim to achieve the purpose of hand pose estimation by using fewer parameters than the state-of-theart model. In our model, a variety of operations to reduce the number of parameters are used, finally resulting in the number of parameters being reduced to 1.5M. This is 21% less than the state-of-the-art model. Finally, we test our model on the common hand pose estimation dataset RHD. From the experimental results, considering the accuracy and model parameters, we have found that our method can reach 0.9713 on PCK@0.2, and outperform the current model with parameters less than 10M.

Multimodal Medical Image Fusion Based on Hybrid Bilateral Filter and Contrast Adjustment Model

Yingmei Zhang, Hyo Jong Lee
Department of Computer Science and Engineering, Jeonbuk National University, Jeonju, Korea

Abstract—The purpose of multimodal medical image fusion (MMIF) is to integrate images of different modes with different details into a result image with rich information, which is convenient for doctors to accurately diagnose and treat the diseased tissues of patients. Encouraged by this aim, this paper proposes a novel method based on a hybrid bilateral filter (HBF) and contrast adjustment model. First, HBF is applied to decomposing the input images to obtain the structure layer and energy layer, which have the property of detail preservation. Then, two fusion rules based on structure tensor operator (STO) and contrast adjustment model are designed, which greatly improve the image performance from the perspective of balancing between information retention and enhancing contrast. Experiments demonstrate that the proposed method has superior performance compared with the state-of-the-art fusion methods.

Age-Invariant Face Recognition Methods: A Review

Kedimotse Baruni, Nthabiseng Mokoena, Murali Veeraragoo, Ross Holder
Defense and security, Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa

Abstract— Face recognition is one of the biometric technologies that is mostly used in surveillance and law enforcement for identification and verification. However, face recognition remains a challenge in verifying and identifying individuals due to significant facial appearance discrepancy caused by age progression. Especially in applications that verify individuals from their passports, driving licenses and finding missing children after decades. The most critical step in Age-Invariant Face Recognition (AIFR) is extracting rich discriminative age-invariant features for face recognition and minimizing age-related features. The variation of facial appearance across aging can be solved using three methods, namely, generative (aging simulation), discriminative (feature-based) and deep neural networks methods. This work reviews and compares the state-of-art AIFR methods to address the work that has been done to extract rich discriminative age-invariant features by maximizing identityrelated features and minimizing the age-related features from facial images of individuals (subjects) across aging. According to the literature reviewed, it has been shown that deep neural networks and discriminative methods perform well in AIFR and deep neural network AIFR methods outperform discriminative AIRF methods.

Speeding-Up the Particle Filter Algorithm for Tracking Multiple Targets Using CUDA Programming

Jinhua Zhang, Wenkai Guan, Cristinel Ababei, Henry Medeiros, Richard J. Povinelli
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI, USA

Abstract—Object detection and tracking are essential tasks in many computer vision applications. One of the most popular tracking algorithms is the particle filter, which is widely used for real-time object tracking in live video streams. While very popular, the particle filter algorithm suffers from increased computational runtimes for high-resolution frames and large numbers of particles. In this paper, we investigate the use of CUDA programming as a method to parallelize portions of the particle filter algorithm in order to speed-up its execution time on compute systems that are equipped with NVIDIA GPUs. Experiments that compare a CPU sequential version, as the base case, with the CUDA parallelized version demonstrate an achievable speed-up of up to 7.5x for a 3840x2160 video resolution, and 9216 particles on a computer equipped with an NVIDIA Tesla K40c GPU.

3D Face Reconstruction from Front and Profile Images for Low Computational Devices

Sankarshan Dasgupta, Mehdi R. Zargham
University of Dayton, Dayton, Ohio, USA

Abstract— Three dimensionality (3D) face re-construction is an advanced and challenging feature for computer vision, and our vision is to implement it using various methods to bring 3D models closer to reality. Although many algorithms for construction of 3D model from two dimensional (2D) images are present, we propose a new approach using front and one profile image with various image processing techniques for small computing devices. Basic methods are utilized for generation of the UV-map of texture, but as its core element, it relies on the Haar Cascade face detection algorithm. For structure or mesh, a shape detector with 68 landmarks is implemented to identify the shape of the face in the image and compare it with our own dataset for most similar structure for faster 3D model construction. Though we have achieved good results from the proposed approach, there is potential to improve by making the model an identical replica.

An Efficient Fingerprint Enhancement Algorithm for Access Control Identification

Cecilia N. Mphahlele, Topside E. Mathonsi
Department of Information Technology, Tshwane University of Technology, Pretoria, South Africa

Abstract — Fingerprint Enhancement Algorithm for Access Control Identification, is an emerging system that provides good and efficient service for access thought out., it plays a very important feature in accessing any access anywhere in the world. There are so many issues around it and methods implemented but failed because of the interruptions on the system which are due to the over ink, dry skin, cut, bruises, and fingerprint misalignments. The main objective of this research is to be able to identify the great performance of an automatic fingerprint identification or to make sure that the verification system will be of good quality for the input fingerprint images. This focuses on fingerprint recognition and how such a system would be implemented. During the verification process, results are to be retrieved based on the identification which is provided and the matching which is performed. It is very essential to help in incorporating the fingerprint enhancement algorithm in the details extraction module. The whole improvement process involves using algorithms that will overcome any possible limitations that are found in the biometric systems. In this research, the fast fingerprint enhancement algorithm is going to be presented, which can very easily to the improvement of the clarity of the ridges and again the valley structures to the input fingerprint images which are based solely on the frequency and local ridge orientation. For fingerprint verification and identifications, we have many approached which are available to be used. There is certainly one method, that represents the fingerprint by its local features, for example, terminations and bifurcations, the method is based on minutia. The simulation tools which will be used in this study will help to improve the quality of inaccurate and askew fingerprint image during the authentication.

Image Segmentation and Classification: Alien Plant Detection Using MATLAB

Ibidun Christiana Obagbuwa, Kutlo Baldwin Mogorosi, Tshepang Duncan Morapedi
Department of Computer Science and Information Technology, Sol Plaatjie University, Kimberley, South Africa

Abstract - Biological diversity is threatened by invasive alien plant species (IAPs) that have spread outside of their natural range. They are not natural to the ecosystem and may cause economic or ecological harm. They have a detrimental influence on biodiversity, resulting in the decrease or extinction of native species because of water competition and disturbance of local ecosystems and ecological processes. IAPs have harmed natural biodiversity in nearly every ecosystem type on the planet and are one of the most serious threats to biodiversity. Traditional methods used to identify and detect invasive alien plants such as utilizing direct sampling field-based techniques or making visual estimation have provided average success. These methods are prone to errors, time-consuming and labour-intensive. Remote sensing techniques offer a concise swift approach for detecting and mapping invasive alien plants. However, remote sensing hardware technology such light detection and ranging (LIDAR) are expensive. The emergence of computer vision and machine learning has provided an inexpensive alternative which can be deployed from a mobile device. This paper illustrates the application of computer vision in the detection of invasive alien plants in South Africa.

Face Recognition Using MATLAB

Ibidun Christiana Obagbuwa, Stefany Bam, Dineo Tiffany Moroka
Department of Computer Science and Information Technology, Sol Plaatie University, Kimberley, South Africa

Abstract - This work sought to investigate how face recognition can be implemented in MATLAB to correctly detect and identify an individual using their face. Face recognition is a biometric technology that is used to recognize and authenticate a detected face in images or videos. This technology can be used in various industries for various purposes. The main goal of this work is to correctly authenticate an individual face using Convolutional Neural Networks (called AlexNet) in MATLAB.

Combining Techonologies for Aiding Search Missions with Drones

Fernando G. Tinetti, Oscar C. Valderrama Riveros
III-LIDI, Fac. de Informatica, UNLP - CIC Provincia de Bs. As, La Plata, Argentina;
III-LIDI, Fac. de Informatica, UNLP - Universidad Cooperativa de Colombia, Ibague, Colombia

Abstract—This paper presents the combination of technologies in an application for aiding search missions. Unmanned Aerial Vehicle (UAV), wireless communications, and image processing and classification technologies are combined for generating warnings to an operator in a search mission. Thus, the UAV would be used to cover a specific ground area communicating images the system should be able to classify. In case a positive image is classified (i.e., one with the object/s being searched for), an operator is warned so that the final decision is made by an expert. In this way, the mostly bored task of looking at images without object being looked for is avoided, and the expert attention could be focused on images with a high probability of having the object/s being looked for. A cloud service is used for image classification (IBM Watson) along with a previously proposed UAV system architecture and a low-cost drone SDK. A proof of concept is described, giving implementation details, and preliminary results. Obtained results show that the general system is a viable solution and set a starting point for enhancing several subsystems for better performance.

Partial Attention CenterNet for Bottom-Up Human Pose Estimation

Jiahua Wu, Hyo Jong Lee
Division of Computer Science and Engineering, Jeonbuk National University, Jeonju-si, Republic of Korea

Abstract—The typical bottom-up human pose estimation methods can be divided into two steps, keypoint detection and grouping. The traditional keypoint regression-based methods exploit an effective backbone (like HRNet) and different prediction heads to acquire the body center and body joint. Then they utilize the offset between the body center and body joint to figure out the grouping task. In this paper, we first propose a body branch module and keypoint attention module to improve keypoint detection and keypoint regression. In body branch module, we exploit a multi-branch structure for keypoint detection and keypoint regression. Each branch represents a part of human body. In keypoint attention module, two simple yet reliable pooling layers are adopted to extract the attention areas of different kinds of keypoints. Combining these two modules, we propose a Partial Attention CenterNet for multi-person human pose estimation. The proposed method outperforms the traditional keypoint regression-based methods. Experiments have demonstrated the obvious performance improvements on COCO dataset brought by the introduced components.

A Study of DQN using VisionTransformer as an Image Extractor

Toshki Hatano, Toi Tsuneda, Satoshi Yamane, ...
Kanazawa University, Kanazawa, Japan

Abstract—In recent years, the field of image recognition has seen rapid technological innovation, and in addition to the CNN-based models used in the past, models using transformers have also been devised. However, there are no results on the verification of DQN using the latest image recognition models in existing research. In this study, we examined the performance of DQN using VisionTransformer as an image extractor, which has become a hot topic in recent years. The performance was not stable due to problems caused by the different architecture from the resulting CNN.

Low-Resolution Face Recognition on Multi-Person Indoor Environment Using Convolutional Neural Network

Greg C. Lee, Yu Che Lee, Cheng-Chieh Chiang
Department of Computer Science and Information Engineering, National Taiwan Normal University, Taipei, Taiwan;
Department of Information Technology, Takming University of Science and Technology, Taipei, Taiwan

Abstract—Face recognition has been widely applied in many systems in our life. These applications have reached good accuracies of face recognition when face images can be captured in a good condition, such as enough resolution. However, in an indoor environment, a surveillance camera can often cover a wide area with multiple persons; that means only lower resolutions of face images are available in face recognition. This paper presents a face recognition approach for low resolution images using convolutional neural network (CNN) on multi-person indoor environment. Our methods first detect face regions by the YOLOv3 approach and then recognize face images by the trained CNN model. Experiments are performed in an indoor classroom to capture face images with resolutions 20x20 to 70x70. Moreover, face images are with different time to test the stability of our proposed face recognition over months.

Determining Performance of Computer Vision Models using Generalizability, Robustness and Elasticity Score

Aishwarye Omer, Hien Nguyen
Mathematics and Computer Science Department, Pennsylvania State University, Harrisburg, Pennsylvania, USA

Abstract—Performance measurement of computer vision models provides information about their ability to classify objects. However, their performance gets affected in the real-world environment. We propose a modification for the metric called Generalizability, Robustness, and Elasticity score (GRE), which is used to determine the efficiency of the computer vision models. Specifically, we use unaltered Visual Question Answering (VQA) datasets and develop three new datasets for each attribute of the GRE score. The new datasets pass through three novel serial processes designed to enhance the quality of the datasets. The new datasets have a better distribution of feature information of the objects in the original dataset. Their performance is measured by running the datasets on three models specifically modified for our experiment. Two out of three models perform better on our new datasets and provide a better GRE score. We prove that our system works and can provide better results than the conventional method of measuring the performance of computer vision models.

A Symmetric Nonnegative Matrix Factorization Approach for Image Reconstruction

Yueyang Wang, Bahram Shafai
Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts, USA

Abstract—Nonnegative Matrix Factorization (NMF) has been proved to be a powerful method in data processing, and also shown success in applications such as feature extraction and image representation. In this paper, we propose two symmetric matrix based method, Symcom and Symize, to achieve square strategy (SQR) in SQR-NMF. This integration process allows the matrix to preserve symmetry property associated with images to enhance the image reconstruction. Simulation results show that Symcom performs better on super wide or super long data matrices and Symize achieve better results on symmetrical data matrices.

Fusion of Visual and Textual Features for Table Header Detection in Handwritten Text Images

Addisson Salazar, Jose Ramon Prieto, Enrique Vidal, Gonzalo Safont, Luis Vergara
Institute of Telecommunications and Multimedia Applications, iTEAM, Universitat Politecnica de Valencia, Valencia, Spain;
Pattern Recognition and Human Language Technology, PRHLT, Universitat Politecnica de Valencia, Valencia, Spain

Abstract— This paper introduces a new procedure to improve table header detection in handwritten text images from the fusion of the posterior probabilities provided by two baseline classifiers. Each classifier considers a different modality, namely visual or textual features. Both baseline classifiers implements convolutional neural networks, particularly adopting the U-Net architecture. Four fusion methods are considered: the mean; linear discriminant analysis and random forest as meta-classifiers; and a recently developed method called alpha integration. The testing dataset consisted of 89 page images drawn from the Passau dataset. The improved performance provided by the fusion methods in the specific experiments is interesting considering the complexity of the challenging problem approached. In terms of area under the receiver operating characteristic curve the best results were obtained by alpha integration. This method incorporates least mean square parameter optimization. The improvement is relevant in the context of the targeted problem.

Artificial Intelligence in Support of Welfare Monitoring of Dairy Cattle: A Systematic Literature Review

Lucas Mendes Lima, Victor Calebe Cavalcante, Mariana Guimaraes de Sousa, Claudio Afonso Fleury, Diogo Oliveira, Eduardo Noronha de Andrade Freitas
Federal Institute of Goias, IFG, Goiania-GO, Brazil;
Florida State University, FSU, Florida, USA

Abstract— Context: Although agribusiness corresponded to more than 20% of Brazil's Gross Domestic Product (GDP), most livestock is under manual control and manual monitoring. Additionally, alternative technologies are either uncomfortable and stressful, or expensive. Now, despite the great scientific advances in the area, there is still a pressing need for an automated robust, inexpensive and (sub)optimal technology to monitor animal behavior in a costeffective, contact-less and stress-free fashion. Overall, this niche can leverage the benefits of Deep Learning schemes.

High-density 3D Reconstruction in a Large Space Using Single Camera and 2D LiDAR

Sangha Oh, Soon-Yong Park
School of Electronic and Electrical Engineering, Kyungpook National University, Daegu, South Korea

Abstract— We propose a method of 3D reconstructing a large indoor space using a 2D LiDAR and a single camera. This device that combines a 2D LiDAR with a single camera scans 360° areas using a rotating stage. We use this device to scan indoor buildings, solve resolution constraints for rotational stages, and show highdensity 3D reconstruction results. It also introduces how to perform indoor reconstruction in a wide area by combining scanned information from various locations.

Optimal Bilateral Filtering of CT Images

Ivo Draganov, Veska Gancheva
Radiocommunications and Videotechnologies Department, Technical University of Sofia, Sofia, Bulgaria;
Programming and Computer Technologies Department, Technical University of Sofia, Sofia, Bulgaria

Abstract—In this paper a general optimization algorithm is proposed for tuning the parameters of a bilateral filter when processing Computed Tomography images, containing Additive White Gaussian Noise. The Peak Signal to Noise Ratio and the Structural Similarity Index Measure are the target parameters during the optimization with the explicit aim of achieving their maximums. The operation of the optimal configuration of the bilateral filter is compared with the results from filtering of the same images with the Gaussian and average filters. Positive results are obtained and the proposed optimization is considered applicable not only for Computed Tomography images, but also for Magnetic Resonance Imaging, multispectral and hyperspectral images, etc.

Metricam: Fast and Reliable Social Distancing Analysis in Online Security Cameras

Bruno Padilha, Pedro Losco Takecian, Decio Lauro Soares, Andre Luis Schwerz,
Leonardo Tadashi Kamaura, Joao Eduardo Ferreira
Department of Computer Science, University of Sao Paulo, Sao Paulo, Brazil;
Department of Computing, Federal University of Technology (UTFPR), Campo Mourao, Paran, Brazil

Abstract—Distance measurements taken from 2D camera images are subject to the correct estimation of the camera's perspective, that is, the spatial mapping from 2D points imaged by a camera to the correspondent 3D ones in the real world. Current solutions to solve this 3D reconstruction are either dependent on the estimation of vanishing points through the detection of straight lines on targeted images or by employing sophisticated sensors and deep learning algorithms, which require expensive training on huge annotated datasets. Nevertheless, none of those approaches provide the required level of precision and accuracy for social distancing evaluation. In this paper we present Metricam, a real-time lightweight software system for security cameras that computes a 2D to 3D mapping using computational geometry and uses the DBSCAN clustering algorithm to evaluate social distancing evaluation. With Metricam, we have been able to identify several places prone to agglomeration inside the Butanta˜ campus of the University of Sao Paulo, and provide the local ˜ authorities with valuable information to fight off the pandemic.

Simultaneous Analysis of fMRI and EEG Biosignals: a Multimodal Fusion Approach

Luis Manuel Pereira, Addisson Salazar, Luis Vergara
Institute of Telecommunications and Multimedia Applications, Universitat Politecnica de Valencia, Valencia, Spain

Abstract—This paper presents a proposal of new analyses for data from functional magnetic resonance images and electroencephalographic signals acquired simultaneously. Considering the current state of the art in this field, the methodology is proposed in the context of multimodal fusion that can be applied in early and/or late stages of the processing. Several problems such as spatial and time synchronization of the data and possible solutions to deal with them based on over-sampling or under-sampling are discussed. The principal objective of this ongoing research consists of increasing temporal and spatial resolution for recognition of activation zones of the brain (zones of interest) during cognitive tasks. Some preliminary results of 3D reconstruction of the brain volume from electroencephalographic signals of a subject carrying out an oddball task are included.

Weight Perception Analysis using Electroencephalographic Signals

Murtaza Aslam, Fozia Rajbdad, Omer M. Soysal
Electrical and Computer Engineering Department, Louisiana State University, Baton Rouge, Louisiana , USA;
Computer Science Department, Southeastern Louisiana University, Hammond, Louisiana, USA

Abstract— Haptics is the study of coupling human sense of touch and motion with computer systems. In this study, we analyze various weight perceptions represented by electroencephalographic brain signals; a pressure tactile is used as a stimulus. The state of the weight perception is represented by relative power of alpha and beta sub-bands of subjects. Independent Component Analysis is applied for post-processing. In the final stage, Linear Discriminant Analysis and Support Vector Machine techniques are utilized for classification of the brain signals as either lifting a weight or being in resting state. Our proposed method achieved satisfactory performance, when Support Vector Machine with Radial Basis Function kernel classifier and relative power features which are obtained from alpha sub-band of four EEG channels are used.

Improving classification of Composite Materials Measured by Ultrasounds using Classifier Combination

Addisson Salazar, Alberto Rodriguez, Luis Vergara
Institute of Telecommunications and Multimedia Applications, Universitat Politecnica de Valencia, Valencia, Spain;
Universidad Miguel Hernandez de Elche, Departamento de Ingenieria de Comunicaciones, Elche, Spain

Abstract—This paper presents a procedure to classify composite materials based on a combination of the output from several single classifiers. The goal is to distinguish between polyester based resin materials doped with graphene nanoparticles from homogenous materials. This is interesting for material quality control in several industries. Two different concentrations of dopant were tested. The materials were measured by nondestructive testing ultrasounds using through-transmission. Time, frequency and statistical features are extracted from the ultrasound signals. Physical elastic constants are also considered. We demonstrate that optimal fusion based on alpha integration of the posterior probabilities from single classifiers improve accuracy and variability of the classification.

Late Fusion for Improving Intrusion Detection in a Network Traffic dataset

Addisson Salazar, Luis Vergara, Gonzalo Safont
Institute of Telecommunications and Multimedia Applications, Universitat Politecnica de Valencia, Valencia, Spain

Abstract—This paper presents a method for an intrusion detection system based on late fusion of classifiers. The proposed method was tested in the analysis of a network traffic dataset considering up to 14 classes of anomalous traffic. Several data quality issues were solved by preprocessing: missing data, nonnumeric data types, and imbalance of the data classes. The high dimension of the data was reduced by a feature selection method. The results demonstrate the capabilities of late fusion to improve classification accuracy and stability of the intrusion traffic detection from the ones obtained by the individual classifiers.

Corrupting Data to Remove Deceptive Perturbation: Using Preprocessing Method to Improve System Robustness

Hieu Le, Hans Walker, Dung Tran, Peter Chin
Boston Unversity, Boston, Massachusetts, USA;
Microsoft Research, USA

Abstract—Although deep neural networks have achieved great performance on classification tasks, recent studies showed that well trained networks can be fooled by adding subtle noises. This paper introduces a new approach to improve neural network robustness by applying the recovery process on top of the naturally trained classifier. In this approach, images will be intentionally corrupted by some significant operator and then be recovered before passing through the classifiers. SARGAN - an extension on Generative Adversarial Networks (GAN) is capable of denoising radar signals. This paper will show that SARGAN can also recover corrupted images by removing the adversarial effects. Our results show that this approach does improve the performance of naturally trained networks.

A Discussion of Optimization about Stereo Image Depth Estimation Based on Multi-baseline Trinocular Camera Model

Hanrong Wang, Ming Li, Jie Wang, Yang Li, Sidan Du
School of Electronic Science and Engineering, Nanjing University, Nanjing, China

Abstract—The huge computational complexity and occlusion problems make stereo matching a big challenge. In this work, we use multi-baseline trinocular camera model to accelerate the stereo matching algorithms and improve the accuracy of disparity estimation. We propose a special scheme named the trinocular flexible disparity searching range (FDSR) to accelerate the stereo matching algorithms. In this scheme, we optimize stereo matching by reduce disparity searching range. Based on FDSR, we proposed the FDSR-MCCNN for trinocular stereo matching. According to the evaluation results, the FDSR-MCCNN could not only reduce the computational complexity but also improving the accuracy. Moreover, the optimization schemes we designed can be extended to the other stereo matching algorithms which possesses pixel-wise matching cost calculation and aggregation steps. We proved that the proposed optimization methods for the trinocular stereo matching are effective and the trinocular stereo matching is useful for either improving accuracy or reducing computational complexity.

CSCI-ISMC:
MOBILE COMPUTING, WIRELESS NETWORKS, & SECURITY

Mimic: A Remote Webcam Device Over WebRTC

Logan Crandall, Mason Roberts, En Cheng
Department of Computer Science, The University of Akron, Akron, Ohio, USA

Abstract— Whether it be classes, work, or social gatherings, the global pandemic has shown that the world can operate in a virtual space. The growing popularity of working from home saw a steady rise in video conferencing. In a virtual space, it is often necessary that participants have a webcam to have a meaningful connection with others. However, some participants may not own a webcam. Mimic is a software solution to allow the use of a webcam-enabled device, such as a smartphone, as a webcam on a computer that otherwise does not have a camera. The goal of Mimic is to provide the ability to use a secondary device as a webcam input without the need to install any additional software on that mobile device. Mimic can accomplish this by building a web client that leverages the power of modern web browsers and WebRTC to stream the webcam video feed from a mobile device to another computer.

WPA3: The Greatest Security Protocol That May Never Be

Glen Sagers
Computer Science and Information Systems, Southern Utah University, Cedar City, Utah, USA

Abstract—Wi-Fi Protected Access version 3 (WPA3) is the newest security standard for wireless networks. Ratified in 2018, and mandatory for devices bearing the Wi-Fi trademark since July of 2020, the protocol has many security improvements over previous versions. It has better encryption and key sharing than the older WPA2 protocol. Unfortunately, adoption of WPA3 is likely to be very slow, just like its predecessors. These delays have nothing to do with the protocol, and everything to do with human factors and legacy systems. Many users do not understand either why they need new security measures, or how to implement them. Legacy systems, specifically Internet of Things (IoT) devices which can only connect to WPA2 networks, are widespread, and probably will not be updated. This paper is a call for industry awareness and action.

Web-based Malware Detection for Android OS

Hani Alshahrani, Ali Alshehri, Abdulrahman Alzahrani, Huirong Fu
College of Computer Science and Information Systems, Najran University, Najran, Saudi Arabia;
University of Tabuk, Tabuk, Saudi Arabia;
University of Hafer Al Batin, Hafer AlBatin, Saudi Arabia;
Department of Computer Science and Engineering, Oakland University, Michigan, USA

Abstract—Smartphones and tablets have become some of the most consumed electronic devices because they revolutionized many aspects of our lives. Android is one of the most popular mobile operating systems (OS) used by mobile users. It captures more than 74 percent of the market share in 2020. These devices store vast amounts of valuable data ranging from personal information (e.g., text messages and contacts) to company information for when companies apply bring your own device (BYOD) policy. Due to its popularity and stored information these devices have been targeted by cybercriminals which introduce a lot of issues to users such as stealing private information, make calls or send SMS messages to premium numbers etc.. In this paper we introduce a public web-service tool to analyze Android applications. It runs inspected applications on real Android devices and provide users with a real time analysis report. Moreover, other researchers can utilize our analysis results to learn more about malicious applications behavior to develop a robust technique to detect these type of applications.

Optimized Edge, Fog and Cloud Computing Method for Mobile Ad-hoc Networks

Josip Balen, Davor Damjanovic, Petar Maric, Kresimir Vdovjak
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia

Abstract—The future of mobile systems relies on big data and fast information processing to improve efficiency, user experience, and system autonomy. With the massive growth of intelligent and mobile devices, along with the development of computational and communication technologies, a vast amount of data is being generated which needs to be processed and distributed as fast as possible, while reducing costs. Therefore, gathering and data processing, as well as efficient distribution of messages across the network, is the key problem of future mobile systems. In this paper, we are proposing a cost-effective method that relies on fog and edge computing principles called FOGO (FOG Optimization). The method proposes a three layer architecture that utilizes an upgraded existing infrastructure where the mobile nodes will create a fog service layer that could provide processing of medium-sized messages and the distribution of the results across the network. Furthermore, cloud computing is also considered but only in a case of necessity for a large-sized data processing. The proposed system architecture is described along with the method flowcharts, operating algorithms, and proposed message structure. Furthermore, the metrics for the system performance evaluation are proposed, as well as possible application domains.

Simulating Realistic IoT Network Traffic Using Similarity-based DSE

Peter Brand, Joachim Falk, Tanja Maier, Jurgen Teich
Friedrich-Alexander-Universitat Erlangen-Nurnberg, Germany

Abstract—In the research for novel network technologies and protocols, network simulators play a crucial role by offering a fast and safe evaluation environment. Here, it is crucial that the simulated traffic is similar to representative and realistic network traffic. Such traffic depends on all connected devices and their application behaviors. These behaviors are modeled by simulation parameters inducing a vast design space—comprising all possible parameter settings that each may lead to a different simulated traffic trace. To tackle the vast parameter design space, we propose a general methodology that efficiently explores this design space to find optimal parameter settings that result in simulated traffic traces that are most similar to a given trace. For an Internet of Things case study and two state-of-the-art similarity measures, we show that we can improve trace similarity by up to a factor of 184 and 4.92, respectively.

An Enhanced VoIP Codec Transcoder to Enhance VoIP Quality for IP Telephone Infrastructure

Thubelihle S. Zulu, Topside E. Mathonsi
Department of Information Technology, Tshwane University of Technology, Pretoria, South Africa

Abstract— Poor VoIP quality in IP telephone infrastructure is a major concern and it can affect business growth, especially to businesses that deals with interacting with a client over the phone. Speech or audio signals are usually affected by codec mismatch, packet loss, and jitter which affect user perception of voice quality. VoIP telephone system is growing at a rapid speed and has received much attention because of its call cost internationally and national and fewer resources needed compare to traditional voice telephone systems or public switched telephone networks. The main aim of this study is to develop a solution that will provide an enhanced voice quality in VoIP platform systems by implementing the amended VoIP codec transcoding system that auto negotiates VoIP codec with the intention of preventing VoIP codec mismatch via standalone and software VoIP codec transcoding system. An experimental research with technological tools such as SIP (Session Initiation Protocol) phone, asterisk PBX (Private Branch Exchange) systems and SBC (Session Border Control) will be conducted. A practical test will be carried out in any working environment with the converged network in order to test results or findings to solve the problem of codec mismatch with the intention of enhancing Voice quality and avoid calls dropping issues in IP telephone infrastructure. This study is introducing an amended VoIP codec transcoding system that auto-negotiate VoIP codec in order to prevent codec mismatch and enhance voice quality hence codec mismatch is not only the major concern for VoIP quality, VoIP quality can be affected by many factors, such as packet loss, jitter, packet delay, and bandwidth but this study is focusing on the codec mismatch.

(n, n) XOR-based Visual Cryptography Schemes with Different Meaningful Shares

An-Hui Lo, Justie Su-Tzu Juan
Department of Computer Science & Information Engineering, National Chi Nan University, Nantou, Taiwan

Abstract—Visual cryptography (VC) is a branch of cryptography, and its most feature is that it can be decrypted using only human vision. With the reduction of hardware costs, XORbased visual cryptography is the latest development of VC, which can solve the problem of low image quality of the restored image in OR-based VC. In this paper, we will give three XOR-based VC scheme with meaningful shares which can import n different cover images. Experimental results and analysis show that our schemes are better than currently known schemes.

An ML Based Anomaly Detection System in Real-Time Data Streams

Javier Jose Diaz Rivera, Talha Ahmed Khan, Waleed Akbar, Muhammad Afaq, Wang-Cheol Song
Department of Electronic Engineering, Jeju National University, Jeju, Republic of Korea;
Department of Computer Engineering, Jeju National University, Jeju, Republic of Korea

Abstract— Due to the advancements in machine learning and artificial intelligence applied fields, network anomaly detection systems have experienced an evolution from traditional signaturebased methods for intrusion detection. Nonetheless, as security measures evolve also more sophisticated attacks are constantly being developed by hackers. The importance of a robust anomaly detection algorithm is evident, but also, a real-time data feeding mechanism for minimizing the reaction-time impact is required. Moreover, DDoS attacks can flood the network data channels with more than thousands of packets per second with the latent effect of overloading most traditional monitoring systems that rely on data storage. Due to this, the research presented in this paper focuses its efforts on implementing a real-time data streaming system for network anomaly detection that can operate during a high volume of traffic data. The solution includes the deployment of a flow collector platform connected to Apache Kafka for receiving NetFlow data from network switches. Also, real-time big data processing techniques are applied through Apache Spark, where the ML anomaly detection is triggered. The detection of anomalies is performed by a combination of the unsupervised learning clustering algorithm k-means and the supervised learning classifier KNN (k- nearest neighbors). Finally, a monitoring system consisting of an ELK stack collects historical data for further evolution of the ML algorithms.

DCLPP: A Distributed and Cooperative Approach based on a Local Path Planning for Multi-sensors Patrolling - Application to Rapid Bushfire Detection

Elie Tagne Fute, Doris-Kholer Nyabeye Pangop, Emmanuel Tonye
Department of Computer Science, University of Dschang, Cameroon;
Department of Computer Engineering, University of Buea, Cameroon;
Department of Electrical and Electronics Engineering, University of Yaounde I, Cameroon

Abstract - This study aims to show that rapid detection of fire risk areas in a forest can be done by keeping the forest under constant and regular surveillance by a small number of mobile wireless sensors. The proposed surveillance approach is based on continuous local path planning. We propose a genetic algorithm suitable for planning a longer path, allowing the sensor to visit a cluster of cells while traversing as many cells as possible in the cluster. The evaluation of the proposed approach shows its effectiveness in reducing environmental idleness and in quickly detecting emerging events.

Key Factors Affecting Mobile App Usage Intensity During Isolation: A Cross-National Study

Hyeon Jo
IT Laboratory, Realsecu, Busan, South Korea

Abstract—The emergence of COVID-19 has had a huge impact on the world, and citizens are feeling extreme anxiety and fear. Governments around the world are implementing various social measures such as social distancing and mask wearing to prevent the spread. University students have not been able to attend faceto-face lectures and are still receiving education through online lectures. Due to reduced outdoor activity time and restrictions on private gatherings, students communicate socially or seek psychological stability through smartphones. This study identifies the antecedent factors that affect the usage intensity of mobile app use during the isolation period. Data were collected from 360 university students in South Korea and Vietnam. PLS-SEM (partial least squares structural equation modeling) method was conducted. The findings of this study reveals that the usage intensity of smartphones positively influence the usage intensity of mobile apps. The results show that affective risk perception has a significant effect on usage intensity of smartphones. Cognitive risk perception is significantly associated with the usage intensity of mobile apps. Regulatory environment significantly affects both the usage intensity of mobile apps and smartphones. Several implications were described for researchers and practitioners.

Distributed Algorithms for k-Coverage in Mobile Sensor Networks

Rinku Sen, Rajib Kumar Das
Department of Computer Science and Engineering, University of Calcutta, Kolkata, India

Abstract—A critical problem in mobile sensor networks is how to move the sensors so that the intended area is k-covered. There are two approaches to solve the problem. Base station assisted centralized algorithm or a distributed algorithm where each sensor determines its next movement based on local information (position of nearby sensors). We have first proposed a distributed algorithm for 1-coverage. Here, a sensor computes the force acting on itself by its neighbor at a distance d. This force is repulsive/attractive depending on d being less/more than the desired distance. For k-coverage, we have reduced the desired distance by $\sqrt{k}$. We have run a set of experiments to evaluate the performance of the proposed distributed algorithms.

Evaluation of the Security Level of Aruba Network Wireless Networks in Educational Institutions

De la Cruz De la Cruz Paola, Dominguez Dominguez Rusbel
Universidad de Montemorelos, Mexico

Abstract— Wireless networks have made it possible to connect mobile devices, providing great advantages for today's communication. However, this has made communications a target of continuous security attacks. The risks can range from being left without connectivity in an organization, to the theft or hijacking of valuable information for institutions. To minimize the risk of a cyber-attack on a wireless network, there are complex and external methodologies to evaluate security, detect failures in the network and help find solutions. This research proposes the use of a methodology based on "ethical hacking" in order to evaluate, analyze and expose the vulnerabilities that may have a "cluster" formed by Aruba Network equipment within an institution of higher education in northeastern Mexico. In the application of this methodology a total of forty-seven vulnerabilities were found, four of them considered critical, eighteen are high risk, twenty-four have a medium score and one vulnerability is in the low risk level.

A High-Performance 5G/6G Infrastructure for Augmented, Virtual, and Extended Reality

Jeffrey Wallace, Angelica Valdivia
Rocket Technology Systems, LLC, Washington, DC, USA

Abstract—Fully developed 5G application delivery services at the mobile edge with reliable "round trip" low latency such as the "motion to photon" measurements in augmented, virtual, and extended reality (AR/VR/XR) systems are not commercially available yet. In fact, these features may well be a part of 6G infrastructures since viable business models, applications, and supporting technology such as military tactical headsets capable of AR/VR/XR in day/night operations are just coming to market and becoming available. Additionally, the first generations of AR/VR devices and 5G infrastructure have provided lessonslearned for technology developers eager to field useful systems with customer appeal. This paper describes the infrastructure for one of the Department of Defense 5G programs focused on AR/VR for military training.

ACLSGO: Ant Colony Link State Aware Geographic Opportunistic Routing Protocol for VANET

Fatima Belamri, Samra Boulefekhar, Djamil Aissani
Research Unit "LaMOS Modeling and Optimization of Systems", Bejaia University Targa Ouzemmour, Algeria

Abstract—The unique characteristics of the Ad Hoc Vehicle Network (VANET), such as high mobility and dynamic network topology, greatly affect data transmission. Indeed, selecting a promising route that forward data is subject to multiple Quality of Service (QoS) constraint such as link failure. In this paper, we propose the Ant Colony Link state aware geographic Opportunistic routing protocol ACLSGO. Based on path robustness and Ant Colony Optimization (ACO), the proposed ACLSGO selects the best candidate node set and determines the optimal priority node to transmit the data. Simulation results show that ACLSGO protocol outperforms LSGO in terms of packet delivery ratio, throughput and also the average end-to-end delay is considerably reduced.

CSCI-ISSE:
SOFTWARE ENGINEERING

Converting Android Native Apps to Flutter Cross-Platform Apps

Yoonsik Cheon, Carlos Chavez
Department of Computer Science, The University of Texas at El Paso, El Paso, Texas, USA

Abstract — Flutter is a development framework for building applications for mobile, web, and desktop platforms from a single codebase. Since its first official release by Google less than a few years ago, it is gaining so much popularity among mobile application developers, even being regarded as a game-changer. There are, however, millions of existing native apps in use that meet the requirements of a particular operating system by using its SDK. Thus, one natural question to ask is about rewriting an existing native app in Flutter. In this paper, we look at the technical side of this question by considering Android apps written in Java. In a small case study, we create a Flutter version of our existing Android app written in Java to support both Android and iOS by rewriting the entire app in Flutter. We share our development experience by discussing technical issues, problems, and challenges associated with such a rewriting effort. We describe our approach as well as the lessons that we learned.

Crossed Analysis of Three-variable for Early Pre-diagnosis of COVID-19

Ana Maria Campos Mora, Diego Sanchez Buitrago, Leonardo Juan Ramirez Lopez
Engineering Faculty, Universidad Militar Nueva Granada, Bogota, Colombia

Abstract— This research study presents the creation of a method of analysis of three physiological variables considered vital for the early diagnosis of COVID-19: Body Temperature, Heart Rate, and Blood Saturation. The applied method was the cross-analysis of variables to obtain triage-type criteria for classifying the individual in one of the three states: Prevention (yellow), Warning (Orange), and Alarm (Red) for each particular case. As a result, an automatic analysis algorithm was developed to support the physician in preventive treatment. It is possible to generate the warning states and classify the situation when making a report according to its condition by validating the results. The algorithms are published on Github to make them available to the scientific community in general and thus solve the early diagnosis.

Grammar-based Fuzzing Tool Using Markov Chain Model to Generate New Fuzzing Inputs

Hamad Al Salem, Jia Song
Computer Science Department, University of Idaho, Moscow, Idaho, USA;
Najran University, Najran, Saudi Arabia

Abstract—Nowadays, Fuzzing has been one of the promising techniques in software testing field. It supports testers and software developers to find bugs and issues in their applications. Fuzzing is automated hence it saves time and energy for the testers. However, the biggest drawback of fuzzing is that it usually cannot reach the deeper levels in an application because randomly generated fuzzing inputs are most likely not satisfying the format requirement of the application. To solve this problem, we conducted our previous research and developed a grammarbased fuzzing tool by extracting grammars from sample inputs. Then, we take the sample input files and modify them to generate effective fuzzing input files. By modifying the sample input files, the generated fuzzing input can test deeper code because of meeting the format requirement of the program. However, the limitation of modifying sample input files is that the tool is limited by the quality of the sample input file. In addition, the tool uses the same commands order which prevents going to other locations in the target program. In this paper we propose our new technique of generating fuzzing input files with the help of Markov chain model on top of our previous work. The new tool learns the commands order, the probability of occurrence of each command, and the structure of sample input file. It utilizes the analyzed grammars to generate completely new fuzzing files. We have tested the DARPA CGC dataset and the new tool has successfully crashed another 7 programs that were not discovered by our previous work.

A Comprehensive Framework for Measuring and Improving API Usability

Jeff Tian, Abdullah Bokhary, Sultan Alanazy
Department of Computer Science, Southern Methodist University, Dallas, Texas, USA;
Department of Software Engineering, College of Computer Science and Engineering, University of Jeddah, Saudi Arabia;
College of Computer Science and IT, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Abstract - In this paper, we develop a comprehensive framework for measuring and improving API usability, consisting of both a comprehensive measurement framework and an analysis/modeling framework. Our measurement framework consists of both a set of direct measurements of API usability and a set of indirect measurements of factors that may influence API usability. Our analysis/modeling framework validates the measurement framework empirically and establishes predictive relations between its two sets of measurements. Such predictive relations can help us improve API usability through recommended actions applied on its influence factors. Examples from a preliminary case study are included to demonstrate the viability and potential effectiveness of our approach.

Real Time Issues of Validation and Verification in Software Development Cycle

Evan Yousif, Dafer Alali, Mohammed Mahmoud, Mohamed Zohdy
Department of Computer Science and Informatics, Oakland University, Michigan, USA;
Department Computer & Electrical Engineering, Oakland University, Michigan, USA

Abstract - In the recent, rapid advancements of the internet, the software development life cycle becomes an essential process in ensuring that the created software systems are errors free. During the software development cycle, several steps are followed for the growth of efficient systems. Validation and verification processes are employed in software development to guarantee the systems are of the best quality and compatible, as well as to confirm the built software meets the client's specifications, requirements and needs. The validation and verification processes are a combination of various analyzing and testing activities done during the software development cycle. These processes are effected with multiple issues that could prevent them from delivering the best in terms of identifying the defects in the product. This leads to the production of poor-quality software that does not fully meet the client.s functional and non-functional needs. The aim of this research is to obtain the main causes of failures of the validation and verification processes, analyze them and suggest the best effective curative that would alleviate these faults. This research involves system developers and experts who greatly contribute to identifying the issues affecting validation and verification. Results associated with advancement in technology, problems of faulty tools, inappropriate environment, human aspects and lack of use of validation and verification procedures in the early stages of system development will be presented. Some of the proposed solutions are; involving the validation and verification experts in early stages, the use of formal methods in validation and verification, providing experts with mathematical and scientific skills and employing appropriate testing tools and environment. This paper also highlights the survey questions and how the analysis was done for the response to the rationale and need for the research.

Design of Software Applications using Access and Actions Control Policies based on Trust

Marcel Danilescu, Laura Danilescu
ASWIC Ltd., Chisinau, Republic of Maldova;
Department of Computers and Information Technology, "Dunarea de Jos" University of Galati, Galati, Romania

Abstract— Granting access to an organization's information resources is an issue that is the subject of numerous research works, with different approaches. This paper addresses access and action control policies based on the levels of trust given to them. The internal organization of an enterprise implies the existence of a hierarchy of departments, structured in a tree, in which data and information are circulated both ascending and descending. The staff of the organization is the one who carries out various processes, which consist of actions, workflows and/or information flows and events. In order to participate in these processes, a certain level of trust is assigned to the person. The association between the level of trust given to a person and the value of trust attributed to an object leads to the generation of policies implemented by computer applications that use access and actions control based on trust. The creation of these policies and their updating is done from a Policy Creation Point. A Policy Storage point is used to store all policies. The Document Status Point is the location where the document status matrix is located. Thus, The Document Storage Point is the space where documents are stored in electronic format. By creating them, a single point of access to policies is established for their creation and updating, a point where policies are stored, a point of storage of the workflow applied to documents and the active process, and a point of documents storage.

Towards a Framework to compare Blockchain Platforms for Smart Contracts

Giuseppe A. Di Lucca, Maria Tortorella
Department of Engineering, University of Sannio, Benevento, Italy

Abstract— Blockchain technology represents an innovation from both a structural and application point of view. Since its start it has undergone a strong and fast development and, thanks to the implementation of smart contracts, it has been applied in different application contexts, such as business, healthcare, manufacturing, IoT applications and much more. The interest around blockchain has brought to the definition of several blockchain platforms facilitating their development and application. Thus, to choose the blockchain platform that is more suitable to support a specific business need is a strategic problem. This paper proposes an evaluation framework, i.e. a set of quality attributes, to characterize and compare the different blockchain platforms helping to identify the most suitable one.

Preservation of Manual Changes and Provenance for Data Quality using the Nano Version Control Repo

Lukasz Machowski, Tshilidzi Marwala
Synthesis Software Technologies, Johannesburg, South Africa;
Department of Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg, South Africa

Abstract— A new data structure called the Nano Version Control (NanoVC) repo emerges from computer science and the software industry. This data structure efficiently encodes entities at the nano-scale of the modelling spectrum and stores the provenance for that entity. The repo provides an intuitive representation of the history and data-lineage of the entity. Some provenance information can be computed on demand because of the repo structure. A simple algorithm for preservation of manual changes in the light of new data and changing algorithms utilizes the commit history in the repo to give us a sustainable way to merge information while keeping the provenance intact.

Nano Version Control and "Robots of Robots" - Data Driven, Regenerative Production Code

Lukasz Machowski, Tshilidzi Marwala
Synthesis Software Technologies, Johannesburg, South Africa;
Department of Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg, South Africa

Abstract— The Corona pandemic highlights the need for sustainable production systems using automation. Traditional automation is fragile and hard to get right. This method converts one hard problem of producing sustainable production code into three simpler problems: data, patterns and working prototypes. By using agent-based simulation and NanoVC repos for agent arbitration, we create a simulated environment for the approach. Patterns developed by people are used to transform working prototypes into templates where data is fed through to create the robots that create the production code. Having two layers of robots allow early implementation choices to be replaced. Robot of Robots encode a legacy of the person that designed them, thus reducing the fragility of the production code.

Regression Testing of Mobile Apps

Zeinab Abdalla, Anneliese Andrews, Ahmed Alhaddad
Department of Computer Science, University of Denver, Denver, Colorado, USA

Abstract - Because mobile applications, or apps, are becoming essential in our personal lives and at work, and mobile applications update frequently, it is important that developers perform regression testing to ensure their quality. In this paper we adapt FSMWeb approach for selective regression testing to allow for selective regression testing of mobile applications. We apply rules to classify the original set of tests of mobile app into obsolete, retestable, and reusable tests based on the types of changes to the model of mobile Apps. New tests are added to cover portions that have not been tested.

Automatic Test Cases Generation for C Written Programs Using Model Checking

Daniset Gonzalez Lima, Raul E. Gonzalez, Pedro Mejia Alvarez
Computer Science, CINVESTAV unidad Guadalajara, Jalisco, Mexico

Abstract—The present work focuses on the development of a tool that automatically generates coverage criteria based test cases from a C written program. For accomplishing this, the tool translates the C code into PROMELA and generates specifications based on the wanted coverage criteria. Once the model (PROMELA code) and specifications are obtained, it uses SPIN model checker for executing the verification and generating counterexamples which can be used as test cases.

Virtual Basketball Training Platform

Tyson Howard, Zizhong J. Wang
Virginia Wesleyan University, Virginia Beach, Virginia, USA

Abstract - The goal for this project is to create a simple and efficient online platform for beginner basketball players who have a passion for learning the fundamentals of basketball. The system website gives users the ability to learn basic skills like dribbling, shooting, passing, and defense. An admins page is created so the administrator can check on the players' progress and answer any questions they may have. All of these functions were created by using HTML, PHP, CSS, BOOTSTRAP, and MYSQL.

Modern HCI for Mobile Applications, Study and Challenges

Ali Abbiss, Brent Soloniewicz, Cameron Jordan, Chandler Connolly, Chris Wiegand,
Corey Dutkiewicz, Mohammed Mahmoud
Department of Computer Science and Engineering, Oakland University, Rochester, Michigan, USA

Abstract—In the past few decades, we have witnessed the rise of the mobile computer, and with that rapid development comes new ways humans are able to interact with these devices. With the multitude of embedded sensors included in these smart devices, software tools can be developed that can mimic, and even replace, the functionality of traditional physical tools such as maps using mobile Global Positioning Systems (GPS), or digital cameras. The goal of this paper is to research and discuss our findings on how Human Computer Interaction (HCI) with mobile devices differ from traditional HCI, how these differences can be used for new HCI opportunities, as well as the ways our interaction changes in response to a new modern HCI.

A Survey on Security and Human Related Challenges in Agile Software Deployment

Neha Singh, Palak Patel, Soma Datta
University of Houston - Clear Lake, Houston, Texas, USA

Abstract – Over the last two decades, agile methods have transformed and brought unique changes to software development practice by strongly emphasizing team collaboration, customer involvement, and change tolerance. However, it also has some challenges. This study aims to accumulate these challenges for a large-scale agile development. It has found there were 11 human related challenges and solution for scaling agile methodology. In addition, there were other significant challenges on agile security, and organization.

A Metric for Software Service Outsourcing Contracts within the Scope of the Brazilian Federal Public Administration

Washington Almeida, Felipe Furtado, Fernando Escobar
Center of Advanced Studies and Systems of Recife - C.E.S.A.R., Recife, Brazil;
University of Minho - UMinho, Guimaraes, Portugal

Abstract— The Brazilian public administration outsources software development and has a large budget for hiring companies for this purpose - called software factories. However, as they use public resources, these contracts are subject to control agencies' audits, like the Federal Court of Accounts (TCU). There are several metrics for measuring software and for measuring software development and improvement projects. However, the TCU prohibited the use of the Technical Service Unit (TSU) metric due to the lack of standardization and, as a result, the damage that this brings to the comparison of budget estimates between public contracts - a practice in the processes of public procurement in Brazil. Within this context, this work presents the results of adopting a process model to create a metric applied to service outsourcing contracts, based on a historical basis, expert opinion, and complying with current Brazilian legislation. The process was ap-plied in a federal judiciary court, resulting in a contract in the information technology area, with 12 professional profiles involved in activities to support the software engineering processes, demanded through a pre-defined catalog of ser-vices. Alternative metrics evaluated: TSU, outsourced government personnel (i.e., body shop), and the costs of expanding the number of public employees to per-form the service. In the analyzed context, the TSU proved to be the best alternative financially and in its applicability, despite not being standardized. Also, the work concluded that when a metric is based on a historical basis, technical and methodological supported, and adapted to the context, the risk of malpractice is mitigated, observing positive results emerging from the applied innovation; moreover, considering the importance of the adaptation to the context, most of the metrics contracted by the government would not be comparable in terms of budget estimates.

CSCI-ISSC:
Smart Cities and Smart Mobility

Enhanced Cybersecurity for Safe Smart World

ChoongHee Han
Security Information Strategy Office, Korea Power Exchange, Naju-si Jeollanam-do, Republic of Korea

Abstract— Unfortunately, convenience is getting more overemphasized, the risk of smart world is getting more neglected. Smart worlds like smart cities, smart homes, smart factories, smart traffics are making web-based IT systems more and more. Many research papers tells us that web-based IT systems are fundamentally vulnerable. Truly, it is very difficult to defend against every cyber attack. Definitely, it is impossible to think about safe smart world without cybersecurity. It is really needed to reduce the risk of smart world. If access from overseas is not necessary, blocking cyber threats from abroad is the best way to reduce the risk of cyber infringements for smart world.

IoT Smart Home Devices' Security, Privacy, and Firmware Labeling System

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

Abstract—Smart home is one of the most popular Internet of Things (IoT) implementations. Smart home devices are appliances and devices that are interconnected through the internet, allowing the homeowner to remotely control and monitor the home. Equipped with smart home devices, a smart home provides homeowners with convenience and even cost savings. Since smart home devices are collecting data about the home and are connected to the internet, their security becomes an important issue. However, due to the lack of a security standard for smart home devices, manufacturers implement different security methods in their products. When purchasing a smart home device, there is no obvious way to find out how secure the device is. Therefore, this research aims to provide an IoT devices' labeling system that covers the security, privacy, and firmware factors of the IoT smart home devices. The proposed label is expected to help and educate IoT smart home consumers and make aware of the potential issues associated with smart home devices, hence protect them from being a victim of many problems and attacks when using the devices. This labeling system is designed specifically for smart home devices selling on the Saudi Arabian market.

Autonomous Cycle of Data Analysis Tasks for Scheduling the Use of Controllable Load Appliances using Renewable Energy

Jose Aguilar, Juan Giraldo, Manuela Zapata, Andres Jaramillo, Luis Zuluaga, Maria D. R-Moreno
Universidad de Alcala, Escuela Politecnica Superior, ISG, Alcala de Henares, Spain;
CEMISI, Universidad de Los Andes, Merida, Venezuela;
GIDITIC, Universidad EAFIT, Medellin, Colombia

Abstract - With the arrival of smart edifications with renewable energy generation capacities, new possibilities for optimizing the use of the energy network appear. In particular, this work defines a system that automatically generates hours of use of the controllable load appliances (washing machine, dishwasher, etc.) within these edifications, in such a way that the use of renewable energy is maximized. To achieve this, we are based on the hypothesis that depending on the climate, a prediction can be made of how much energy will be generated and, according to the behavior of the users, the energy demand required by these appliances. Following this hypothesis, we build an autonomous cycle of data analysis tasks composed of three tasks, two tasks for estimating the required load (demand) and the renewable energy produced (supply), coupled with a scheduling task to generate the plans of use of appliances. The results indicate that it is possible to carry out optimal scheduling of the use of appliances, but that they depend on the quality of the predictions of supply and demand.

An Efficient CNN-based Approach for Automatic Brazilian License Plate Recognition

Joao Pedro Cabral, Vitor Gaboardi Dos Santos, Camila Souza, Alison Silva, Abraao Dantas, Nelio Cacho, Frederico Lopes, Daniel Araujo
Digital Metropolis Institute, Federal University of Rio Grande do Norte, Natal, Brazil;
Department of Informatics and Applied Math, Federal University of Rio Grande do Norte, Natal, Brazil

Abstract—Automatic License Plate Recognition (ALPR) is increasingly becoming the target of many studies in computer vision in recent years due to its great applicability in urban environments. In Brazil, a new type of license plate called Mercosul went into circulation in the past year, and few works proposed to include the detection of this kind of plate. To address this issue, we trained our models with approximately sixty-five thousand sample images from the Federal Highway Police (PRF) captured by the institution's radars that were not previously labeled. Using three Convolutional Neural Networks (CNN) it was possible to create an ALPR system with an accuracy of 97.91% in the UFPR-ALPR Dataset, 95% in the SSIG SegPlate Dataset and 79.77% in the UFRN Dataset presented in this paper. These results outperform the most reliable studies for Brazilian license plates.

A Lightweight Ontology for Real Time Semantic Correlation of Situation Awareness Data Generated for First Responders

Iosif Angelidis, Elena Politi, George Vafeiadis, Danai Vergeti, Dimitris Ntalaperas, Nikos Papageorgopoulos
UBITECH, Athens, Greece

Abstract—In case of an emergency, the immediate action of the first responders is crucial for saving human lives. Their intervention requires specialized instrumentation, available at all times and easily accessible, which meets stringent requirements in terms of detection accuracy, quick localization, and reduction of false alarms. This work proposes a novel ontology-based methodology which integrates data from IoT devices in the frame of a Situation Awareness (SA) semantic model. The proposed model aims at providing the conceptual representation of core entities, which will be represented by concepts and will cover specific aspects of the SA domain, such as proper decision making during the course of an emergency operation, conceptual representation of critical information required for such tasks and of important information flows which potentially exist between involved actors. Finally, we set the scene towards validating the efficiency and efficacy of our proposed model directly in the field, through seven (7) use cases.

Model-Based Testing of Smart Home Systems Using EFSM and CEFSM

Afnan Albahli, Anneliese Andrews
Department of Computer Science, University of Denver, Colorado, USA

Abstract—The Internet of Things (IoT) has become a hot software technology. The number of devices connected to the Internet has been growing dramatically and is expected to continue to grow. This increase causes a huge challenge for software quality. New testing approaches need to be developed and investigated to assure quality and efficiency of such systems. The primary challenge with IoT devices is that the functionalities can be highly variable due to the device type and the way these devices are connected. In order to maintain feasibility of these functionality systems need to be modeled. This paper proposes a testing approach for a smart home system (SHS) modeled by Extended Finite State Machines (EFSMs) and Communicating Extended Finite State Machines (CEFSMs). We generate tests for individual devices in the SHS as well as the interaction between devices.

On Cognitive Management and Non-Causal Reasoning for Enabling Highly Automated Driving

Ilias E. Panagiotopoulos, Konstantina Karathanasopoulou, George J. Dimitrakopoulos
Department of Informatics and Telematics, Harokopio University of Athens, Greece

Abstract—Highly Automated Vehicles (HAVs) have become a trend, and also a hotspot of research in recent years, aiming to support, or even to replace, human drivers. Their goal is mainly to strengthen the driver's sensing ability and to reduce the control efforts of the vehicle itself. Moreover, on-board communications equipment helps vehicles to have a model of their complex driving environment that includes the presence (meta-knowledge) of other entities sharing the same driving scene. Therefore, cognitive decisions should be taken in an automated manner, being able to operate HAVs each time in the best available Level of Autonomy (LoA), by responding quickly not only to causal reasoning effects, which depend on present and past inputs from the external driving environment, but also to non-causal reasoning situations, which depend on future states associated with the external driving scene. The present study aims to tackle exactly this challenge by introducing an on-board cognitive decisionmaking functionality, which operates on the basis of collecting information from various sources, intelligently processing it, integrating knowledge and experience and, finally, selecting the optimal LoA.

Drowsiness Detection Systems: Comparison and Technical Criteria for Industrial Deployment

Ralf Luis de Moura, Debora de Souza Martins, Luciana de Land Farias Ceotto
Operational Technology Architecture, Vale S/A, Vitoria, Brazil;
Operational Technology Architecture, Tata Consultancy Service, Vitoria, Brazil

Abstract—Avoiding accidents at work has been a major concern for companies for years. Their impacts, such as financial, image, reputation, and human losses, are less tolerated worldwide. Traffic accidents in motor vehicles correspond to many accidents at work and can be caused by several factors. One of the main factors recurrently associated with accidents is when the driver sleeps. When sleeping, the driver completely loses control over the vehicle's actions, which can cause serious accidents. One of the approaches to prevent the driver from falling asleep is to predict the state of sleep (drowsiness). There are several ways to drowsiness detection, such as physiological data readings, driver's images, and blink frequency. There is no international standard on this topic or standout technology, so choosing the right market solution is problematic. This study analyzes these solutions, proposes selection criteria with functional and nonfunctional requirements, evaluates meaningful market solutions based on these criteria, and offers recommendations for deployment following the best technological practices.

Risk Assessment in the Context of Dynamic Reconfiguration of Level of Driving Automation in Highly Automated Vehicles

Ilias E. Panagiotopoulos, Konstantina N. Karathanasopoulou, George J. Dimitrakopoulos
Department of Informatics and Telematics, Harokopio University of Athens, Greece

Abstract—Advanced Driver Assistance Systems (ADAS) constitute a field that continues to attract immense research BY promising significant advantages and novelties to the manner in which we drive vehicles, facilitating several of the driver's operations and the passengers' journey, as well as protecting the vehicle from undesired situations. With the advent of Automated Vehicles (AVs), the research and development in ADAS will be intensified, so as to holistically undertake the responsibility of getting a vehicle safely from one point to another point. Risk assessment and reliability analysis are a cornerstone of the evaluation of ADAS and their probability of success in completing a prescribed mission of an AV. In this paper, the classical Failure Mode and Effect Analysis (FMEA) technique is applied to investigate the risk assessment regarding real time adaptation of the Level of Driving Automation (LoDA) in AVs. This analysis is crucial as high risky events are evolved in the transition mode of LoDA related to hardware, sensors, software failures, front obstacles or crashes, dense traffic congestion, adverse weather or road conditions, etc. Through this analysis, an efficient approach is developed by exploring the reliability of the LoDA transition in an AV operation and its impact on the behaviour of both the AV and the driver.

Optimal Operation of a Microgrid for Energy-aware Buildings by Decision-based Algorithm

Teodora Mindra, Gheorghe Florea, Oana Chenaru, Radu Dobrescu, Lucian Toma
Societatea Inginerie Sisteme, Bucharest, Romania;
Faculty of Automatic Control and Computers, University Politehnica of Bucharest, Romania;
Faculty of Power Engineering, University Politehnica of Bucharest, Romania

Abstract— Microgrids are becoming a popular approach for increasing the efficiency in use of renewable resources. This paper presents a microgrid control and optimization structure adapted to the specific requirements of an office building in Romania. A model is used to evaluate process behavior and maximize the impact in energy efficiency. The results of the proposed optimization strategy emphasize the advantages of using microgrids for increasing energy efficiency. This is achieved by upgrading smartness of buildings, and also by using a decisionbased optimization algorithm for real-time power balancing.

P2P Power Trading between Nanogrid Clusters Exploiting Electric Vehicles and Renewable Energy Sources

Sangkeum Lee, Hojun Jin, Sarvar Hussain Nengroo, Dongsoo Har
Environment ICT Research Section, Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea;
The Cho Chun Shik Graduate School of Green Transportation, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea

Abstract— P2P energy trading addresses direct energy exchange between peers, thereby energy from small-scale distributed energy resources in households, workplaces, factories, and other locations is exchanged among neighborhood energy prosumers and consumers. A novel method for real-time P2P power trading between nanogrid clusters based on cooperative game theory is proposed in this paper. Cooperative P2P power trading is used as a powerful aid for the power management of nanogrid cluster involving electric vehicles and renewable energy sources wind turbine and photovoltaic energy system. For the power management of nanogrid clusters, multiobjective optimization making use of relevant information obtained from the internet of things and from the time-varying production of hybrid wind power and PV power is carried out. As a result, cooperative P2P power trading between nanogrid clusters can save the electricity cost and amount of energy supplied from the grid, as compared to the stand-alone nanogrid clusters without P2P energy trading.

CSCI-ISPD:
PARALLEL AND DISTRIBUTED COMPUTING

Starvation Avoidance Task Scheduling Algorithm for Heterogeneous Computing Systems

Amjad Gawanmeh, Wathiq Mansoor, Sa'ed Abed, Darin Kablaoui, Hala Al Faisal
Department of Electrical Engineering, College of Engineering and IT, University of Dubai, Dubai, UAE;
College of Petroleum and Engineering, Kuwait University, Kuwait

Abstract—High-performance and cost-effective Heterogeneous Computing Systems (HCS) are highly dependent on the efficiency of task scheduling, which has been proven to be NPcomplete. This makes scheduling problems a vital issue which researchers have attempted optimization in different aspects under the expense of tradeoffs. The problem of scheduling a single Directed Acyclic Graph (DAG) has been extensively researched and studied, however most of the researches in this area did not consider the occurrence of multiple applications represented as multiple DAGs in real case scenarios. To fill that gap, this paper aims at providing a fairness aware, time efficient, (HCS) static scheduling algorithm that meets deadlines, reduces makespan, and avoids starvation while considering the possibility of having to schedule tasks of multiple applications concurrently. Our algorithm, namely, Fair, Deadline-Aware, Time Efficient, Starvation Avoiding (FDATESA) scheduler, relies on four pillars (i) mapping the DAGs into information abundant maps (ii) Task priority heuristic calculation (iii) Assigning tasks to processors' queues (iv) elastic processor time slot calculation for starvation avoidance.

Human-in-the-Loop Automatic Data Migration for a Large Research Computing Data Center

Fang (Cherry) Liu, Michael D. Weiner, Kevin Manalo, Aaron Jezghani, Christopher J. Blanton, Christopher Stone, Kenneth Suda, Nuyun Zhang, Dan Zhou, Mehmet Belgin, Semir Sarajlic, Ruben Lara
Partnership for an Advanced Computing Environment (PACE), Georgia Institute of Technology, Atlanta, Georgia, USA;
Research Computing and Data Facilitation Service, North Carolina State University, Raleigh, North Carolina, USA;
Microsoft Inc., Atlanta, Georgia, USA

Abstract—Most HPC centers face a lack of expertise of data center migrations, as it's a rare event that only a small portion of HPC professionals experience more than once in their entire professional careers. This paper presents how the Georgia Institute of Technology (Georgia Tech) Partnership for an Advanced Computing Environment (PACE) team employed automation to migrate research computing data from the old Rich computing center (Rich) to the new Coda data center (Coda) in 2020. PACE successfully migrated 1844 TB of data for 3550 users without loss of user data. PACE implemented a 'human-in-the-loop' automatic workflow to facilitate the migration, interleaving automated scripts with human-driven reviews, significantly reducing staff time commitment while ensuring the integrity and accuracy of data migrations. PACE deployed a cached data movement strategy which reduced the migration downtime significantly. We share our one-year migration journey for the benefit of the HPC community.

Prediction and Convergence Calculations using Rust-based NAlgebra GLM

Noah W. Scott, Douglas D. Hodson, Richard Dill, Michael R. Grimaila
Air Force Institute of Technology, WPAFB, Wright-Patterson Air Force Base, Ohio, USA

Abstract—Prediction and convergence are techniques used to reduce the network traffic between multiple distributed simulation applications that individually maintain a representation of a "world" that include moving entities. Prediction (often using dead reckoning algorithms) is an approach to estimate the position and orientation of "remote" entities hosted and/or managed by other simulation applications executing within the distributed system. Estimates are made (i.e., calculated) using previously received data, such as velocity and acceleration. As new data is received, a convergence algorithm is often used to update the "remote" entity's position and orientation within the represented world. The term "convergence" is often times referred to as "blending" or "smoothing" as its goal is to avoid visually obvious disjointed "jumps" in movement as updates are received. This work implements the dead reckoning estimation algorithms defined in the IEEE standard for Distributed Interactive Simulation (DIS) in software using the Rust programming language and the NAlgebra GLM package (i.e., crate/library). It also implements a simple convergence algorithm to move entities to their correct locations and orientations. This work is part of a larger design effort to prototype a DIS-compatible interoperability network interface, organized using an Entity-Component-System (ECS).

A Load Balancing Scheme based on Deep Learning in Blockchain Network

Hye-Young Kim, Ji-Hyun Lee
School of Games, Major of Game Software, Hongik University, Sejong-si, South Korea

Abstract— The blockchain technologies taking advantage of decentralization, anonymity, and trust have begun to exert a significant influence on variety applications. A large number of event data will be generated over a period of time in the blockchain networks. There are many different implementations of load balancing with different approaches and purposes in blockchain network. They will generate a large amount of data and resolutions, providing humans with information and the control of events and objects, even in remote physical environments. However, the demands of the interactions in blockchain networks cause heavy traffic or bottlenecks on particular nodes nevertheless decentralized. Therefore, we proposed an agent that applies a deep reinforced learning method to distribute loads in blockchain network. Also, we show the scalability and reliability of the proposed scheme through mathematical modeling.

Non-Blocking Technique for Parallel Algorithms with Global Barrier Synchronization

Arturo Garza, Claudio A. Parra, Isaac D. Scherson
Department of Computer Science, University of California, Irvine, California, USA

Abstract—Sharing data among asynchronous processes is considered to be a hard systems problem when dealing with modern shared-memory multicore systems. Throughout the literature, multiple solutions have been proposed, like the so-called barrier synchronization. A Barrier is a synchronization primitive that provides guarantees that any thread will not continue execution from a given point until all threads have reached that point. This primitive is widely used in different parallel programming models, but it can easily become a hot-spot for performance critical applications due to its global nature as one preempted thread will stop execution of all those other threads waiting at the barrier. This paper suggests a technique to change the global nature of barrier synchronization into a non-blocking synchronization model with lock-free thread progression guarantees. The main idea is to exploit algorithm-based memory access patterns to implement self-synchronizable threads to protect concurrent reads and writes in a shared data structure without explicit use of a barrier primitive. To the best of our knowledge, this is the first attempt to provide a different synchronization mechanism based on the algorithm intrinsic characteristics rather than an explicit use of a global barrier in shared-memory architectures. Our experimental results show factors of performance improvement against its barrier-based algorithm counterpart.

Continuous Variable Quantum Compilation

Federico Rueda, Sonia Lopez Alarcon
Department of Computer Engineering, Rochester Institute of Technology, Rochester, New York, USA

Abstract—In recent years quantum computing has evolved from a research concept subject to reality. Companies and scientists around the world are creating physical quantum systems using different quantum technologies to advance the field. Although most of this focus is on superconducting, qubitbased implementations, recently there has been work in the development of boson-based quantum computers. The basic processing unit for these devices are continuous variable (CV), multi-photonic modes in Gaussian states. Due to the novel nature of this technology there is a lack of work on analyzing the compilation process of current photonic devices. One of the leading companies in this field developed a cross-platform python library for simulating and executing programs on quantum photonic hardware, Strawberry Fields (SF) [4]. This paper will focus on analyzing the SF compilers, and proposing an additional compilation step that will enable future, more flexible hardware implementations.

Distributed Network Optimization for Secure Operation of Interdependent Complex Networks

M. Hadi Amini, Laurent L. Njilla, Ahmed Imteaj, Calvin Mark
Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, Florida, USA; USA Air Force Research Lab, USA

Abstract— The optimal operation of complex networks and critical infrastructures requires solving various large-scale decision-making problems. These problems usually are formulated as optimization problems with several variables and constraints. This leads to the high computational complexity of solving the underlying optimization problem. Hence, we require efficient methods to first model the operational objective function and constraints of the complex networks, and how they can leverage available computational resources to achieve the optimal operation of the entire system. We further need to ensure data security of decision-making entities, e.g., network flow problems, and their impact on the secure operation of the system. The proposed framework and algorithms in this paper include distributed intelligence among heterogeneous agents in a complex network represented by a graph of nodes and edges among them. Our utilized methods act as efficient computational algorithms to solve the underlying optimization problems of these networks in a computationally-efficient fashion. In order to evaluate the introduced distributed algorithm for linear-constrained optimization with a quadratic cost function, we used a random network with different numbers of nodes and edges. We illustrate the run-time and convergence of the distributed method over various networks.

Accelerating Workload Processing with MPI for Porter's Stemming Algorithm

Eduardo Colmenares, Heng Wu
Department of Computer Science, Midwestern State University, Wichita Falls, Texas, USA; Department of Computer Science, West Virginia State University, Institute, West Virginia, USA

Abstract— Stemming is the process of reducing a word to the roots of words known as a lemma. For example, a stemming algorithm reduces the words “Likes”, “likely” to the stem “like”. Stemming is a part of linguistic studies in information retrieval and extraction, it plays a critical role in artificial intelligence (AI), especially in Natural Language Processing. Computers are now equipped with multicore processors and many of the software solutions to a diversity of problem not necessarily take advantage of all these cores to process information and accelerate the processing of workloads. This research presents an effort to accelerate workload processing through the use of Message Passing Interface (MPI) for Porter’s stemming algorithm.

Hierarchical and Density Aware Spectral Clustering Approach to Server Placement for Collaborative Virtual Services

Sakir Yucel
Wexford, PA, USA

Abstract— Boundaries on separate infrastructure domains are blurring as corporations and infrastructure providers are collaborating to offer end-to-end services over networks and infrastructures of various network, service, cloud, content delivery and other infrastructure providers as well as over the customer premises. To capture opportunities and shine in competitive service market, infrastructure and service providers need to excel in addressing the changing customer requirements and in the operations and management of the resources. They should welcome effective collaboration with other network and infrastructure providers in delivering quality services to customers. Effective sharing of the infrastructure resources is essential in meeting the customer demands and reducing the cost. Server placement problem for end-to-end virtual services becomes a crucial optimization challenge for providers in such collaborative environments. In our previous work, we formulated the collaborative virtual server placement problem and suggest density-based clustering algorithms to address this problem. in this work, we hierarchical and density-aware spectral clustering as feasible solutions to the server placement problem in virtual and collaborative service environments.

Parallel Multithreaded Medical Images Filtering

Veska Gancheva
Department of Programming and Computer Technologies, Technical University of Sofia, Sofia, Bulgaria

Abstract—The quality of medical images is paramount. Being of high grade, it guarantees the quality of medical diagnosis, treatment and quality of patient's life through the means of health care or using automate intelligent systems for medical diagnosing, treatment and monitoring. The paper presents the computational challenges in medical images processing. The great challenges are to propose parallel computational models and parallel program implementations based on the algorithms for medical images filtering. Parallel computational model based on two-dimensional filters is designed. The proposed parallel model is verified by multithreaded parallel program implementation. An investigation of the efficiency of medical images filters based on parallel multithreaded program implementation, applying two-dimensional filters on a given list of compressed jpeg medical images and generating output jpeg images for each type of applied filter. The applied filters are Brightness Control, horizontal and vertical filter of Sobel, Laplace and Blur. A number of experiments have been carried out for the case of dataset consisted of 162 whole mount slide images of Breast Cancer (BCa) specimens scanned at 40x and various number of threads. Parallel performance parameters execution time and speedup are estimated experimentally. The performance estimation and scalability analyses show that the suggested model has good scalability.

Out-of-GPU FFT: A case study in GPU prefetching

Peter Pirgov, Lenore Mullin, Rishi Khan
Extreme Scale Solutions, Newark, Delaware, USA;
MoA: Provably Optimal Tensors, Arlington, Virginia, USA

Abstract - In this paper, we propose a decomposition of the N-dimensional FFT and novel transposition strategies to optimize performance for input sizes that do not fit on the GPU. The state-of-theart GPU FFT library, cuFFT, efficiently solves FFT problems that fit in the GPU memory. Additionally, using managed memory, cuFFT can solve problems that exceed the GPU memory, albeit inefficiently due to poor prefetching from the CPU. The major bottleneck in computing the FFT on a GPU is the PCI bandwidth. Therefore, careful prefetching is required to maximize PCI bandwidth. Batches of decomposed input data are sent to and from the GPU to overlap communication with computation. The batches are organized such that the dimension that is stored contiguously is always included to maximize DRAM bandwidth and cache line use. We compare three transposition strategies: CPU based transposition, GPU based transposition, and index-based transposition of the actual FFT and find that GPU based transposition performs the best. Finally, we propose a model that relates the hardware characteristics to the decomposition parameters. We compare our results to the model and to cuFFT on three platforms: a workstation with GeForce GTX 1060, NERSC Cori, and ORNL Summit and show a 2- 3X speedup over cuFFT using managed memory for input sizes that do not fit in the GPU memory.

Odd-Even Flexible Router for High Performance Network-on-Chips

Garrett Hentges, Mostafa Abdelrehim
California State University, Bakersfield, California, USA

Abstract—Recent advancements in fabrication technologies and the ongoing era of deep scale nano transistors extend the umbrella of Moore's law further and further. Manycore architectures with thousands of processing tiles create huge challenges to maintain high communication speeds. From these architectures, Network-on-Chips (NoCs) proves to be the best candidate to achieve these ultimate goals. In this paper, a novel buffering mechanism based on Odd-Even routing algorithm is presented. Old fashioned conventional buffering technique limits incoming packets to be stored in buffers directly connected the input port from which the packets are coming from. Instead, this new mechanism chooses any suitable buffer in the whole router to store incoming packets. In this design we use the buffering restrictions of Odd-Even routing algorithm to prevent any deadlock situations. Our Odd-Even Flexible router successfully outperforms the conventional buffering router.

CSCI-ISCS: COMPUTATIONAL SCIENCE

Scratch-Style Relational Algebra and Calculus

Levi Meston, Jalal Kawash
Department of Computer Science, University of Calgary, Calgary, Canada

Abstract—This paper presents BR∀CE, a Blockly-based, tile-based environment to develop relational algebra and relational calculus queries. The tool runs in a browser, works with any database schema, generates equivalent SQL code to the theoretical query, and more. Specifically, the paper presents the tool's tile-based interface and its compilers. The compilers optimize relational algebra and calculus expressions that correspond to the visual query and translate them to SQL. We present the grammars utilized by these compilers, the abstract syntax trees, and the translation rules. BR∀CE is intended to be an educational tool to tackle the previously reported challenges with teaching and learning theoretical query languages.

An Experimental Evaluation of a Function in Extremal Combinatorics

Kai Wang, Hong Zhang
Department of Computer Science, Georgia Southern University, Statesboro, Georgia, USA;
Department of Computer Science, Georgia Southern University, Savannah, Georgia, USA

Abstract—We investigate the validity of a candidate formula for an extremal function introduced by Ferrara et al. The function is defined to be the minimum degree sum such that every bigraphic pair with a given number of terms in each part and at least this degree sum is guaranteed to have a realization that contains a complete bipartite subgraph of a given size. We show that the formula is valid for some input ranges and invalid for other input ranges. We also show that the difference between the true function value and that given by the formula can be arbitrarily large and conjecture that it may be NP-hard to compute the function value.

Arbitrage Behavior amongst Multiple Cryptocurrency Exchange Markets

Samuel Grone, Weitian Tong, Hayden Wimmer, Yao Xu
Department of Computer Science, Eastern Michigan University, Ypsilanti, Michigan, USA;
Department of Computer Science, Georgia Southern University, Statesboro, Georgia, USA;
Department of Information Technology, Georgia Southern University, Statesboro, Georgia, USA

Abstract—Compared with fiat currencies, cryptocurrencies are usually more vulnerable to speculation and thus lead to massive price fluctuations, which makes exchanging cryptocurrencies a potentially profitable but risky endeavor. We aim to contribute to the understanding of the arbitrage behavior involving multiple cryptocurrency exchange markets. Specifically, we applied a Bellman-Ford based algorithm to detect possible arbitrage opportunities. By investigating historical data from three cryptocurrency exchange markets, i.e., Gemini, Coinbase, and Kraken, we designed experiments to identify how often arbitrage was possible in the past as well as the factors that contribute to the existence of arbitrage. We believe this may bring insights into strategies to stabilize the cryptocurrency exchange markets.

Mechanical Anomaly Detection on an Embedded Microcontroller

Mansoureh Lord, Adam Kaplan
Department of Computer Science, California State University Northridge, Northridge, California, USA

Abstract—This paper explores machine learning on an embedded device to detect anomalies with sophisticated low-power neural networks. We leverage this deep learning approach to detect mechanical anomalies as they occur on a top-load washing machine. We collect normal data from balanced laundry loads and abnormal data from unbalanced laundry loads, as they are being washed by the machine. The normal data is then used to train two different neural network models: autoencoder and variational autoencoder. This model is ported to an Arduino Nano microcontroller mounted to the washing machine. Using the autoencoder model, the microcontroller detects unbalanced washing machine loads with 92% accuracy, 90% precision and 99% recall. The battery life for this autoencoder model is 20 hours on 5 V lithium batteries, which is only 14.9% less than the life of a basic LED-blink application on the same platform.

Ashurbanipal: A Diristry to Document Multimedia Metadata Tools for Transdisciplinary Archives

Anousha Athreya, S. Koby Taswell, Andrew Kang, Ishani Das, Carl Taswell
Brain Health Alliance, Ladera Ranch, California, USA

Abstract—In historical artifact conservation, archiving objects using entity metadata plays a significant role in managing the related versions of the artifacts preserved, recorded and documented at various time points. In this paper, we discuss five fields of study to display the importance of related versions in identifying patterns over time through events in history, cultural heritage, biomedical research, performing arts, and fine arts. We describe our use of the Ashurbanipal diristry to document scholarly research on archiving tools and technologies. We highlight the importance of the provenance infosubset in tracing metadata for cultural objects managed in NPDS repositories and enabling interoperability with existing multimedia bibliographic formats including MARC and BIBFRAME.

A Novel Solution to Perform Real-Time ETL Process based on Non-Intrusive and Reactive Concepts

Flavio de Assis Vilela, Ricardo Rodrigues Ciferri
Department of Computing, Federal Institute of Goias, Brazil;
Department of Computing, Federal University of Sao Carlos, Sao Paulo, Brazil

Abstract—ETL is an essential process required to perform data extraction in knowledge discovery in databases and in data warehousing environments. The ETL process aims to gather data that is available from operational sources, process and store them into an integrated data repository. Also, the ETL process can be performed in a real-time data warehousing environment. This paper present a novel solution to ETL process in a real-time data warehousing environment based on non-intrusive and reactive concept. The Data Magnet has been validated using synthetic data. The results showed the correctiviness of the Data Magnet and the attendance of real-time requirements.

A New Unified Computational Method for Finding Confidence Intervals of Shortest Length and/or Equal Tails under Parametric Uncertainty

Nicholas A. Nechval, Gundars Berzins, Konstantin N. Nechval
BVEF Research Institute, University of Latvia, Riga, Latvia;
Aviation Department, Transport and Telecommunication Institute, Riga, Latvia

Abstract— A confidence interval is a range of values that provides the user with useful information about how accurately a statistic estimates a parameter. In the present paper, a new unified computational method is proposed for simultaneous constructing and comparing confidence intervals of shortest length and equal tails. This unified computation technique provides intervals in several situations that previously required separate analysis using more advanced methods and tables for numerical solutions. In contrast to the Bayesian approach, the proposed approach does not depend on the choice of priors and is a novelty in the theory of statistical decisions. It allows one to exclude nuisance parameters from the problem using the technique of invariant statistical embedding and averaging in terms of pivotal quantities (ISE & APQ) and quantile functions. It should be noted that the well-known classical approach to constructing confidence intervals of the shortest length considers at least three versions of possible solutions and is in need of information about the forms of probability distributions of pivotal quantities in order to determine an adequate version of the correct solution. The proposed technique does not need such information. It automatically recognizes an adequate version of the correct solution. To illustrate this technique, numerical examples are given.

Proposal for the Virtual Library that Enables the User to Control the Bookshelf Based on the Concept of Timeline

Xuezhen Li, Takayuki Fujimoto
Graduate School of Information Sciences and Arts, Toyo University, Japan

Abstract - With daily management and updating of the book classifica-tion, the distribution of book-shelves is changed over time. Through these changes, we can understand the development of knowledge and learning. Therefore, this paper proposes a "dynamic" bookshelf that can be controlled by the specification of timeline settings and popularity beyond the existing Decimal Clas-sification. In addition, "Custom-made" bookshelf can be realized accord-ing to the search demand of each user.

Markov Regime Switching Analysis for the Pandemic and the Dynamics of German Market

Kangrong Tan, Shozo Tokinaga
Faculty of Economics, Kurume University, Kurume, Japan;
Department of Economics, Kyushu University, Fukuoka, Japan

Abstract—This paper deals with the impact of COVID-19 pandemic on the stock market. We focus on the links between the infected cases and the change of the stock market in Germany. We employ the Markov Regime Switching Analysis (MRSA) to expose the situations of COVID-19 pandemic, and the comovement between the pandemic and the market. Through our empirical analysis, we find that MRSA works well to divide the whole time horizon into several intervals with different statistical properties.

AutoTR: Efficient Reformatting Text Spread Out in a DOM Tree for Text-Analytic Applications

Changfeng Yu, Cheng Zhang, Hao Zhang, Jie Wang
University of Massachusetts Lowell, Massachusetts, USA

Abstract—We present an efficient algorithm to automatically reformat text contained in multiple nodes that are spread out in a DOM tree of an HTML file converted from a PDF document. Reformatting text on the fly is needed for hierarchical reading and other text-analytic applications. A naive approach would traverse the DOM tree multiple times, failing to meet the requirement of real-time reformatting. Our algorithm meets the real-time requirement by indexing text nodes and sentences with a pair of tag holders inserted into each text node to allow fast reformatting.

Effectiveness of Privacy Techniques in Smart Metering Systems

Martin Peralta-Peterson, Olivera Kotevska
Electrical and Computer Engineering Department, University of California - San Diego, California, USA;
Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

Abstract—Smart grid technologies enable timely energy billing for residential homes. The ability to react to energy demands during peak hours allows energy providers to conserve power and operate efficiently. However, these data streams are also susceptible to privacy attacks within the energy company and from outside hackers. We implemented four different privacy models: k-anonymous, l-diversity, t-closeness, and ε-differential privacy. We demonstrate the models' effectiveness using a realworld dataset composed of 15 different residential households with energy consumption data spanning over a year.

Robust Asynchronous P Systems for Basic Operations

Naoki Tojima, Akihiro Fujiwara
Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology, Iizuka, Fukuoka, Japan

Abstract—Membrane computing, which is a computational model inspired by living cell activity, has considerable attention as one of new paradigms of computations. Using the membrane computing, several arithmetic operations have been proposed. However, when the cellular activity of an organism is used as computational hardware, the possibility that the cells may be stoped or vanished for some reason cannot be denied. Therefore, a robust algorithm, which continues calculation even if some of cells are stoped or vanished, is nesesarry for membrane computing. In the present paper, we propose robust algorithms for basic operations, which are logic functions and addition, using an asynchronous P system. The proposed P system can obtain a correct solution even if some objects are vanished at any stage of computation.

Cryptocurrency Bounced Back Based on Cryptography Technology during the COVID-19 Pandemic

Fatma Mallouli, Nesrine Khalifi, Aya Hllal, Eman Ferjani, Imen Gmach, Nada Chaabane, Rimah Amami, Rim Amami
Deanship of Preparatory Year and Supporting Studies, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia;
Business Administration Department, The Applied College, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia;
Computer Science Department, The Applied College, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Abstract—How are digital technologies being used to face the pandemic. Several questions may arise when mobilizing digital technologies to respond to the current crisis. Digital technologies may be seen as a gateway to solve many of the problems arising from the crisis: How can we control the spread of COVID-19? How do we continue to provide education to the many people who have to stay at home. At the same time, digital technologies may pose challenges related to several current issue. The covid 19 has accelerated the digitazion of almost all domains that human being uses every day. To ovoid interactions with each other : business, educations , medical sectors and others have adopted various platforms to continue theirs engagement with their colleagues, patients , clients, and students in order to be away from each other. Moreover, Since the pandemic's announcement, house has shifted to have multitask roles. . . the house is no longer the place to rest, it became: the school, the workplace, the entertainment place, and even private clinics. In this research, we analyze the new existing lifestyle, and how technology can fit with these new requirements of our daily life. We introduce the cryptocurrency and its mechanism as ways to support the online transactions within the new corona lifestyle. The aim of this paper is to describe how cryptography reenforced the use of cryptocurrency application distributed. asymmetric cryptography-based security is a provided solution to data-transactions related trust. also, presenting the most popular cryptocurrencies followed by a discussion.

Player Modeling using Behavioral Signals in Competitive Online 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 to match players with similar skills to ensure a satisfying experience for players. In this paper, we focus on the importance of addressing different aspects of playing behavior when modeling players for creating match-ups. To this end, we engineer several behavioral features from a dataset of over 75,000 battle royale matches and create player models based on the retrieved features. We then use the created models to predict ranks for different groups of players in the data. The predicted ranks are compared to those of three popular rating systems. Our results show the superiority of simple behavioral models over mainstream rating systems. Some behavioral features provided accurate predictions for all groups of players while others proved useful for certain groups of players. The results of this study highlight the necessity of considering different aspects of the player's behavior such as goals, strategy, and expertise when making assignments.

Non-Orthogonal HOSVD Applied to 3-D Image Compression

Ionela S. Munayco, Marina Chugunova, Ali Nadim
Institute of Mathematical Sciences, Claremont Graduate University, Claremont, California, USA

Abstract—In this paper we introduce an iterative approach to decomposing tensors into a sum of rank-one non-orthogonal components. The approach uses the repeated application of an idea from the Higher Order Singular Value Decomposition (HOSVD) method that uses power iterations to obtain a rank-one approximation to a tensor. At each step, that rankone approximation is subtracted from the tensor and the process is repeated. We abbreviate our method NOTSVD for Non-Orthogonal Tensor Singular Value Decomposition. For 2- D tensors (i.e., matrices) the process converges to the usual SVD of the matrix and orthogonality is preserved. We present numerical evidence of convergence of the iterative algorithm for approximating 3-D tensors despite the lack of orthogonality. We apply the NOTSVD method to compression of RGB images which can be regarded as 3-D tensors, each color represented by a matrix. We compress the image two ways, one using the matrix SVD applied to each of the three color matrices, and the other applying NOTSVD to the full 3-D tensor. For matrices, we use L 2 norms to characterize differences. We find that the same level of accuracy in reconstructing the image can be achieved with lower storage requirements (higher compression factor) for the NOTSVD method. We also apply the method to a generic 3-D (volumetric) image and show good reconstruction with relatively few components.

Towards Two-Dimensional Game Theory Based Predictive Analytics

Mirhossein Mousavi Karimi, Shahram Rahimi
Department of Computer Science and Engineering, Mississippi State University, Mississippi, USA

Abstract—This article presents a two-dimensional game theory-based model for multi-actor predictive analytics. The model expands the current one-dimensional model developed by BDM to address prediction in more complex problems. So far, the single dimension models have only been able to evaluate the outcome of a problem by looking at the positions of the actors in regards to a single issue. This is while in reality, a player's strategy may depend on the dynamics of multiple issues when interacting with other players. In the first part of this work, the formulation of a two-dimensional model is presented. Then, three case studies have been included to illustrate the efficiency of the proposed tool. To demonstrate our model's ability to reproduce results similar to what reported previously, the first case is considered a one-dimensional issue, investigated in previous articles. The second case deals with a two-dimensional negotiation problem, presented on x- and y-axes, while the third case attempts to predict the oil price on the x-axis considering the renewable energy market share of the players on the y-axis. The presented model is explainable and can be useful to predict the outcome of multilateral issues by taking into consideration current dynamics involving multiple stakeholders with conflicting interests.

Multiple Story Generation by "Story Techniques that are Included in a Story"

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

Abstract—We propose a system that generates stories, using a generation structure. We call this system "a story that includes story techniques." Story techniques are used to generate stories. In the proposed system, the story does include story techniques. The system offers two types of generation: one in which the story is edited from the outside, and another in which the story is edited by a threshold-driven procedure from the inside. The present study simulates story generation, using the proposed system and a story technique known as "coloring."

E-book Circulation System Based on Blockchain

Cheng-Chi Lee, Chung-Wei Li, Chun-Ta Li, Chin-Ling Chen
Library and Information Science, Fu Jen Catholic University, New Taipei City, Taiwan;
Information Management, Tainan University of Technology, Tainan, Taiwan;
Computer Science and Information Engineering, Chaoyang University of Technology, Taichung, Taiwan

Abstract—Along with the trend of the whole society gradually develops in directions like electronic, automatic, etc., and of course library as the representative of culture has no exception, however libraries within this era focuses on the collections and services that most of them base on basic concepts of readers' autonomous operation. Yet they still can't detach from limits of entities such as readers' self-circulation system, electronic collections, etc. We would like to virtualize the whole library that begins with the part of electronic collections. Electronic book, which is often abbreviate to "E-book", can be a better approach to fit the reading hobby of recent society. As we analyse problems that E-book might be faced with are piracy of E-book and cumbersome borrowing process. For these problems, we propose that stores the critical data of Ebooks through Blockchain technology and establishes a private blockchain to operate the whole circulation system. There are some features of this proposal, such as the immutability of blockchain technology is utilised to make high reliability of readers' navigated records which can be traced back, the smart contract is utilised verifying readers in circulation system, etc. In this way, readers can navigate the system by login through browsers instead of borrowing the specific reading device or through the borrowing process, which make readers have a sense like being in the library wherever they are and searching references in the system as soon as they face with problems.

Solving the Dose Distribution Problem in the Radiotherapy Planning Applying the Interior Point Methods

Jackeline del Carmen Huaccha Neyra, Aurelio Ribeiro Leite de Oliveira
Department of Applied Mathematics, University of Campinas, Campinas, Brazil

Abstract—The dose distribution problem in the Radiation Therapy planning consists in optimize the total radiation dosage delivered into the patient. Its main aim is to attack the tumor delivering as radiation as possible, avoiding severe damage the nearby tissues and organs, for that we use a linear programming model to formulate this problem and the Intensity-modulated radiation therapy (IMRT) technique that allows us to deliver nonuniform radiation flow obtaining a best adjustment of radiation doses to the tumor cells. However, in practice, the oncologist can vary the dose according to his/her expertise making it difficult to guarantee an exact quantity of radiation. Therefore in this paper, the doses are considered as triangular fuzzy numbers and we use the surprise function, that can be understood as a penalty for the violation of each constraint, translate these fuzzy constraints in a non-linear function that must be minimized. Thus we present a mathematical model with non-linear and convex objective function with linear constraints for bounding the radiation flow. We propose an algorithm that was designed using a specially tailored Primal-Dual Interior-Point Method to solve this problem. Our algorithm is implemented in MATLAB and tested in real world large-scale problems in which the tumor is localized on the Head and Neck area. The numerical experiments show the amount of radiation delivered in each region of patient and, in general, the developed approach provides favorable solutions for the dose distribution problem.

Clustering Algorithm based on K-value Adaptive Neighborhood Selection

Tianyu Shen, Youwei Wang, Tao Du, Shouning Qu
School of Information Science and Engineering, University of Jinan, Jinan, China; Hoteamsoft, Jinan, China;
School of Information Science and Engineering, Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan, China

Abstract—In this paper, the k-value based adaptive neighborhood selection (KAN) algorithm is proposed, which can adaptively selects the number of K neighbors according to the data density. Then the algorithm calculates the high-density center, expands the clustering according to the k-nearest neighbor similarity, and finally obtains the clustering result. The algorithm effectively overcomes the shortage of setting the number of neighbors in advance based on KNN algorithm, and does not need prior knowledge to set the number of clusters, so as to avoid the wrong results caused by the wrong setting of important parameters. Compared with other algorithms, this algorithm shows obvious advantages.

New Method for the Computation of Generating Functions with Applications

Jun Zhang, Ruzong Fan, Fangyang Shen
Department of Computer Science and Engineering Technology,
University of Maryland Eastern Shore, Princess Anne, Maryland, USA;
Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University, Washington, DC, USA;
Department of Computer Systems Technology, New York City College of Technology,
City University of New York, Brooklyn, New York, USA

Abstract—Generating functions (GFs) are one of the most useful tools for problem solving, as they have been playing an important role in many applications, including but not limited to counting, identity proving, analysis of algorithms, problem representation and solving in combinatorics, etc. The authors have been studying a new transform called Sumudu Transform in a computational approach, in this work, it shall show that Sumudu Transform transfers the exponential generating functions to the ordinary generating functions, and the transform also serves as a new powerful tool in the calculation of generating functions and applications. Applications include new methods in solving differential and integral equations automatically by using generating functions.

Analysis of Routing Entropy of Hyperbolic Trees

Zalan Heszberger, Andras Gulyas, Andras Majdan, Jozsef Biro
Budapest University of Technology and Economics, Budapest, Hungary;
MTA-BME Information Systems Modelling Research Group, Budapest, Hungary

Abstract—Recent results have shown that the memory requirements of destination-based hop-by-hop routing in communication networks can efficiently be estimated by the information theoretic entropy of the forwarding tables placed at the nodes. For calculating and analyzing the memory usage the forwarding tables are to be inferred according to the routing algorithm, then the entropy values can be established. This could be a computationally intensive task, especially in case of large networks operated along complex routing policies making the analysis hard and less tractable. In this paper we focus on a special case, when the routing is based on a spanning tree the so called hyperbolic tree. We show that the routing entropy can efficiently be computed in this case without generating the forwarding tables. Based on this computation, analytical results on routing scalability with respect to memory usage can also be derived, which confirms observations on numerical investigations.

An Algorithm to Solve Systems of Nonlinear Differential-Algebraic Equations with Extraordinary Efficiency Even at High Demanded Precisions

Josef Dobes, David Cerny
Department of Radio Electronics Technicka, Czech Technical University in Prague, Praha, Czech Republic

Abstract—There are more situations when systems of nonlinear differential algebraic equations need to be solved with extraordinary precision. A steady-state analysis (determining the steadystate period of a system after a transient) is a typical case because a vector of unknown variables should be exactly the same after a numerical integration on the period-long interval. Therefore, we need to develop such kinds of numerical algorithms that are computationally effective, even at very high requirements on the accuracy of the results. In the article, an efficient and reliable algorithm for solving systems of algebraic-differential nonlinear equations is characterized first. Unlike usual cases, the procedure is based on a sophisticated arrangement of the Newton interpolation polynomial (i.e., not Lagrange's). This feature provides greater flexibility in rapidly changing interpolation steps and orders during numerical integration. At the end of the paper, two complicated examples are solved to demonstrate that the algorithm's computational requirement is quite low, even at very high demands on the accuracy of the results.

CSCI-ISCB: COMPUTATIONAL BIOLOGY

A Computational Approach to Identify Transcription Factor Binding Sites Containing Spacer Regions

Punithavathi Sundaramurthy, Wendy Lee, Brandon White
Computer Science Department, San Jose State University, California, USA;
Biological Sciences Department, San Jose State University, California, USA

Abstract—A critical challenge in studying gene regulation is deciphering functionally important regions of DNA which when altered, can affect gene activation levels. Bioinformatics tools have been developed to extract motifs from the human genome using methods such as position weight matrices (PWMs), Hidden Markov Models (HMMs), and machine learning (ML). However, these methods are not suitable for motifs with variable spacer regions or when insufficient experimentally validated sequences exist in the literature to build models. In this paper, we present a computational method to identify and extract motifs in conjunction with other high throughput methods such as protein binding microarrays.

Application of Protein Language Models to Low-N Engineering of Metalloproteinase Inhibitors

Elham Khorasani Buxton, Rushabh Patel, Mari R. Toumaian, Maryam Raeeszadeh-Sarmazdeh
Department of Computer Science, University of Illinois at Springfield, Illinois, USA;
Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA;
Chemical and Materials Engineering Department, University of Nevada, Reno, Nevada, USA

Abstract—Targeting metalloproteinases (MPs) has a great potential in developing novel therapeutics as these proteases are known to be responsible in several diseases. Directed evolution was previously used to engineer binding affinity and selectivity of natural inhibitors of MPs, tissue inhibitors of metalloproteinases (TIMPs). However, directed evolution is a time-consuming and laborious process. Machine learning approaches can facilitate directed evolution by narrowing the search for finding optimal mutations. This work investigates the application of state-of-theart pre-trained protein language models in screening MMP inhibitor variants. Features extracted from several protein language models for a low-N experimental library of MMP variants are visualized and compared with regards to their performance in screening strong inhibitors.

Understanding How the Relative Abundance of Candida Species Impacts Transcriptional Regulation in Co-Culture Biofilms

Diksha Kool, Wendy Lee, Katy Kao
Department of Computer Science, San Jose State University, California, USA;
Department of Chemical and Materials Engineering, San Jose State University, California, USA

Abstract—Candida albicans and Candida glabrata are common fungal species that can change from commensal to pathogen due to their ability to form robust biofilms. Candida species are the leading cause of life-threatening conditions like Candidemia, and the existing treatments for biofilm-related infections are suboptimal. Research shows that the relative abundance of the two Candida species promotes biofilm formation, enhances pathogenicity, and increases antibiotic resistance. Thus, focusing on the importance of co-culture, this paper utilizes RNA sequencing to investigate the gene expression leading to biofilm development in co-culture through a time-series study.

SARS-CoV-2 Variant Strains and Viral Phylodynamics of ORF1a-1b Genetic Aspects in South America

Luiz Henrique Pontes Dos Santos, Stela Mirla da Silva Felipe, Cristina Pacheco Santos Martin, Jannison Karlly Cavalcante Ribeiro, Raquel Martins de Freitas, Eric Petterson Viana Pereira, Juliana Osorio Alves, Paula Matias Soares, Valdevane Rocha Araujo, Claudio Gleidiston Lima da Silva, Maria Izabel Florindo Guedes, Vania Marilande Ceccatto
Brazil

Abstract - Bioinformatics tools for online sequence analysis of variants have been used worldwide for the phylogenetic approach of SARS-Cov-2 and their variants. The purpose of this work is to contribute to the settlement of the SARS-CoV-2 genetic of South America pandemic, presenting the ORF1a-1b evaluation. We proposed and reviewed two online bioinformatics pipelines for viral phylodynamic and phylogeographic analysis with an interactive visualization platform. The phylodynamics evaluation of South America shows a strong viral capacity to evoke immunity and an impressive multiplicity of variants in rapid expansion with mutations of potential importance, including ORF 1a-1b. They showed stately vital for infection and lethality, spreading and raising your frequency in South America from 1% to 20-30% in one year of pandemic occurrence.

Visualizing Population Substructures Using Multidimensional Scaling and Data Smoothing

Dimitrios Charalampidis
Department of Electrical and Computer Engineering, University of New Orleans, New Orleans, Louisiana, USA

Abstract—Single Nucleotide Polymorphisms (SNPs) present an important component of a genome's information and have been extensively used in genetics for population structure analysis. SNP data visualization assists in detecting population substructures. However, SNP sequences include thousands or millions of data points. One way to visualize SNP data is through dimensionality reduction. Principal Component Analysis (PCA) has been traditionally used for reducing dimensionality to 2D or 3D with reasonably acceptable outcomes. However, visualizing complex population structures requires more advanced techniques. Recently, t-Distributed Stochastic Neighbor Embedding (t-SNE) has been used for SNP visualization. In this work, a Multidimensional Scaling (MDS)-based method is presented and compared with t-SNE. Although both techniques successfully reveal population substructures in 2D, the MDS-based method better preserves the relative similarity between populations.

Application of Machine Learning to Sleep Stage Classification

Andrew Smith, Hardik Anand, Snezana Milosavljevic, Katherine M. Rentschler, Ana Pocivavsek, Homayoun Valafar
Department of Computer Science and Engineering, University of South Carolina, Columbia, South Carolina, USA;
Department of Pharmacology, Physiology, and Neuroscience,
University of South Carolina School of Medicine, Columbia, South Carolina, USA

Abstract—Sleep studies are imperative to recapitulate phenotypes associated with sleep loss and uncover mechanisms contributing to psychopathology. Most often, investigators manually classify the polysomnography into vigilance states, which is time-consuming, requires extensive training, and is prone to inter-scorer unreliability. While many works have successfully developed automated vigilance state classifiers based on multiple EEG channels, we aim to produce an automated and open-access classifier that can reliably predict vigilance state based on a single cortical EEG from rodents to minimize the disadvantages that accompany tethering small animals via wires to computer programs. Approximately 427 hours of continuously monitored electroencephalogram (EEG), electromyogram (EMG), and activity were labeled by a domain expert out of 571 hours of total data. Here we evaluate the performance of various machine learning techniques on classifying 10-second epochs into one of three discrete classes: paradoxical, slow-wave, or wake. Our investigations have included Decision Trees, Naive Bayes Classifiers, Random Forests, and Artificial Neural Networks. These methodologies have achieved accuracies ranging from approximately 74% to approximately 96%. Most notably, the Random Forest and the Artificial Neural Network both achieve remarkable accuracies of 95.78% and 93.31%, respectively. Here we have shown the potential of various machine learning classifiers to automatically, accurately, and reliably classify vigilance states based on a single EEG reading and a single EMG reading.

Application of Machine learning Ensemble Super Learner for Analysis of the Cytokines Transported by High Density Lipoproteins (HDL) of Smokers and Non-smokers

Seema Singh Saharan, Pankaj Nagar, Kate Townsend Creasy, Eveline O. Stock, James Feng, Mary J. Malloy, John P. Kane
Department of Statistics, University of Rajasthan, Jaipur, India;
University of California Berkeley Extension. California, USA; University of California, San Francisco, California, USA

Abstract - Smoking is a major cause of cardiac and pulmonary disease, cancer, and other inflammation related diseases. Smoking impairs lipid and lipoprotein metabolism. The observed modification and reduction in levels of HDL in smokers has adverse effects on atheroprotective.properties. It has been hypothesized that HDL transports inflammatory cytokines which accelerate tobacco-related diseases. To investigate the role of HDL in the transport of inflammatory cytokines and their detrimental effects on the immune response, it is paramount to compare cytokine levels in HDL for Smoker versus Non-Smoker groups. We isolated HDL from plasma using selected affinity immunosorption of apolipoprotein A-Ibearing lipoproteins, followed by quantitative ELISA of cytokines. We implemented a powerful stacked ensemble Machine Learning algorithm, namely Super Learner with baselearners: Decision Tree classifier, AdaBoost classifier, Bagging classifier, Extra Tree classifier, Logistic Regression and Random Forest classifier. Prediction Accuracy evaluation metric was used to ascertain the separability efficacy of Smoker versus Non-Smoker based on cytokine levels. Super Learner, achieved a 100% prediction accuracy, outperforming all the base learners. Machine learning-enabled Precision Medicine allows the investigation of the role of novel biomarkers such as HDL-transported cytokines which have a potential to generate valuable molecular insights. The discovery that cytokines are transported by HDL presents a new dimension in understanding inflammatory disorders and the potential for therapeutic intervention.

Optimization of a Classical Algorithm for the Alignment of Genomic Sequences with Artificial Bee Colony

Raul Magdaleno Penaloza, Gerardo Reyes Salgado, Andrea Magadan Salazar
Tecnologico Nacional de Mexico/CENIDET, Cuernavaca Morelos, Mexico

Abstract - In this article, the classic algorithms “NeedlemanWunsh” and “Smith-Waterman” are reviewed with which the alignment of genomic sequences is generated; and that of SmithWaterman is improved with the aim of optimizing said processes in order to provide optimal results through the implementation of Artificial Bee Colony. The experiments that are carried out show alternative alignments, to the classical methods, that can be found in the analyzed data.

Improving Gene Expression Prediction of Cancer Data Using Nature Inspired Optimization Algorithms

Payal Patel, Kalpdrum Passi, Chakresh Kumar Jain
Department of Mathematics and Computer Science, Laurentian University, Sudbury, Ontario, Canada;
Department of Biotechnology, Jaypee Institute of Information Technology, Noida, India

Abstract - Cancer being one of the most vital diseases in the medical history needs adequate focus on its causes, symptoms and detection. Various algorithms and software have been designed so far to predict the cancer at cellular level. The most crucial aspect for sorting the cancerous tissues is the classification of such tissues based on the gene expression data. Gene expression data consists of high amount of genetic data as compared to the number of data samples. Thus, sample size and dimensions are a major challenge for researchers. In this work, four different types of cancer microarray datasets are analyzed viz., breast cancer, lung cancer, leukemia and colon cancer. The analysis of the cancer microarray datasets was done using various nature-inspired algorithms like Grasshopper Optimization (GOA), Particle Swarm Optimization (PSO), and Interval Value-based Particle Swarm Optimization (IVPSO). To study the accuracy of the prediction, five different classifiers were used: Random Forest, K-Nearest Neighborhood (KNN), Neural Network, Naïve Bayes and Support Vector Machine (SVM). The Grasshopper Optimization (GOA) outperforms in accuracy compared to the other two optimization algorithms with SVM classifier on leukemia, lung and breast cancer datasets selecting the best genes/attributes to correctly classify the dataset.

On Signaling Dysregulation in Cancer

Youcef Derbal
Ted Rogers School of Information Technology Management, Ryerson University, Toronto, Canada

Abstract— The evolutionary dynamics of cancer underlie its near boundless potential for therapeutic resistance, representing the greatest challenge in the fight against cancer. These dynamics are driven by the evolving diversity of genetic and epigenetic alterations which translate into signaling dysregulations that underpin cancer hallmarks, including the evasion of growth suppression and the reprograming of metabolism. Pathway entropy is proposed as a measure that can differentiate between tumors and normal tissue as well as shed light on acquired therapeutic resistance. The potential utility of the proposed measure is discussed within the context of lung and colorectal cancers.