CRA

Computing Research Association

GRAD Cohort URMD

2019

MARCH 22-23

HAWAII

Underrepresented Minorities & Persons with Disabilities

POSTER SESSIONS

Poster Session

#	Primary Research Area	First Name	Last Name	Current University or Institution	Abstract Title
1	Accessibility	Brenda	Castro	Cornell University	Understanding the Experiences of Teachers of the Visually Impaired with Tactile Materials
2	Algorithms	Sampson	Akwafuo	University of North Texas	A Modified Meta-Heuristics Optimization Model for Post-Disaster Emergency Deliveries with Variable Vehicular Capacities
3	Artificial Intelligence	Gustavo	Aguilar Alas	University of Houston	Dependency-Aware Named Entity Recognition with Relative and Global Attention
4	Artificial Intelligence	Irish	Medina	University of Waterloo	Predicting Short-Term Water Consumption for Multi-Family Residences
5	Artificial Intelligence	Ademola	Okerinde	Kansas State University	Traffic Optimization System using Machine Learning Techniques
6	Bioinformatics	Brienna	Herold	Rochester Institute of Technology	A multi-dimensional method for antigenic characterization of influenza hemagglutinin
7	Bioinformatics	Hassan	Karim	Howard University	Brain computer interfaces (BCI) - Towards an AI framework for safety, privacy and security considering challenges with algorithmic fairness
8	Bioinformatics	Valeria	Velasquez-Zapata	Iowa State University	Next-Generation Screening for Host Interactors of Fungal Effectors in Cereal Immunity
9	Cognitive Science	Lucca	Eloy	University of Colorado, Boulder	Multiparty Regularity in Collaborative Problem Solving
10	Computational Modeling and Simulation	Hirut	Kollech	University of Pittsburgh	Biomechanical Response of the Optic Nerve Head Region in Glaucomatous and Non-Glaucomatous Human Samples
11	Computer Architecture	Hiwot	Kassa	University of Michigan	Fitting Demanding Applications to Heterogeneous Systems
12	Computer Architecture	Maria Isabel	Mera Collantes	New York University	Don't Trust, Verify: A Verifiable Hardware Accelerator for Matrix Multiplication
13	Computer Architecture	Thierry	Tambe	Harvard University	Adaptive Floating-Point Quantization of Deep Neural Networks
14	CS Education	Renesha	Hendrix	Georgia State University	The Rebirth of Engineering Education
15	CS Education	Kenya	Mejia	University of Washington, Seattle	Subject Matter Expertise: Insights on Knowledge Use
16	CS Education	Olivia	Nche	Clemson University	CodeTracesure - A video game to pinpoint student obstacles and facilitate early detection of students who may need help
17	CS Education	JaCoya	Thompson	Northwestern University	Sports Analytics: Learning Data Analytics through the use of Wearable Technologies and Sports
18	Cyber Physical Systems, Internet of things	Amari	Lewis	University of California, Irvine	A Preliminary Design for a Blockchain-Based IoT Paratransit System

Poster Session

#	Primary Research Area	First Name	Last Name	Current University or Institution	Abstract Title
19	Data Mining	Adewale	Obadimu	University of Arkansas at Little Rock	Leveraging Social Media Analytics Tools To Understand Information Operations In Modern Communication Platforms
20	Distributed Computing	Lucianna	Kiffer	Northeastern University	Examining Persistent Hard Forks in Bitcoin and Ethereum
21	HCI	Mahdokht	Afravi	University of Texas at El Paso	Long-Term Multiparty Interaction in Virtual Reality
22	HCI	Oliver	Alonzo	Rochester Institute of Technology	Reading-Assistance Tools for Self-Directed Learning by Deaf and Hard-of-Hearing Computing Workers
23	HCI	Larwan	Berke	Rochester Institute of Technology	Preferences and Requirements of Deaf and Hard-of-Hearing Users for Captions Generated via Automatic Speech Recognition
24	HCI	Tya	Chuanromanee	University of Notre Dame	Evaluation and Comparison of Usability of Four Mobile Breathing Training Visualizations
25	HCI	Steven	Díaz Hernández	University of South Florida	Lower Body Rehabilitation System using Wearable Sensors
26	HCI	Gabriella	Johnson	University of Colorado Boulder	Eye in the Sky: An Audio Description System for Board Gamess
27	HCI	Lindah	Kotut	Virginia Tech	Designing Mobile Technology to Support Hands-Off Interactions
28	HCI	Aboubakar	Mountapmbeme	University of North Texas	Accessible Blockly: A Block-Based Programming Platform for Users With Visual Impairment and Sighted Users.
29	HCI	Larry	Powell	Texas A&M University	Everyday-Inspired Movies: Movie Recommendation System Based on Personal Social Media Posts
30	HCI	Isabel	Rayas	University of Southern California	Affective State Estimation in Human Interactions with Virtual Agents
31	HCI	Wendy	Roldan	University of Washington, Seattle	The Role of Funds of Knowledge in Online Search and Brokering
32	HCI	Jonathan	Saddler	University of Nebraska, Lincoln	Understanding Student Comprehension through Eye Gaze

Poster Session

#	Primary Research Area	First Name	Last Name	Current University or Institution	Abstract Title
33	HCI	Jean	Salac	University of Chicago	An Analysis through an Equity Lens of the Implementation of Computer Science in K-8 Classrooms in a Large, Urban School District
34	HCI	Jomara	Sandbulte	Pennsylvania State University	Exploring Photo Sharing to Engage Intergenerational Families on Health
35	HCI	Ather	Sharif	University of Washington	OptiGain: Automatically Discovering the Optimal Mouse Gain Setting for People With Limited Hand Function through Everyday Computer Use
36	HCI	Mark	Yousef	University of Minnesota	Measuring Psychological Effects of Peer-to-Peer Reputation Systems Involving In-Person Exchanges
37	Internet of Things and Distributed Ledgers	Mayra	Samaniego	University of Saskatchewan	Detecting Suspicious Transactions in IoT Blockchains for Smart Living Spaces
38	Machine Learning	Abiola	Arise	Kansas State University	Identifying Features Important for Predicting Information Diffusion in Twitter Events
39	Machine Learning	Kiante	Brantley	University of Maryland	Non-Monotonic Sequential Text Generation
40	Machine Learning	Adedolapo	Okanlawon	Kansas State University	Cloudbursting: Stay-or-Go?
41	Machine Learning	Deontae	Pharr	Georgia State University	Recurrent Convolutional Neural Networks for Diagnosing Mental Disorders
42	Machine Learning	Angela	Stewart	University of Colorado, Boulder	I Say, You Say, We Say: Using Language to Model Shared Knowledge Construction during Collaborative Problem Solving
43	Natural Language Processing	Angelina	McMillan-Major	University of Washington	Automating Linguistic Annotations in IGT for Low-Resource Languages
44	Natural Language Processing	Lena	Reed	University of California, Santa Cruz	Can Neural Generators for Dialogue Learn Sentence Planning and Discourse Structuring?
45	Natural Language Processing	Andrea	Villanes	North Carolina State University	Epidemiological Disease Surveillance Using Public Media Text Mining
46	Networks	Ranysha	Ware	Carnegie Mellon University	Battle for Bandwidth: Fairness and Heterogenous Congestion Control

Poster Session

#	Primary Research Area	First Name	Last Name	Current University or Institution	Abstract Title
47	Neuromorphic Hardware	Shamma	Nasrin	University of Illinois at Chicago	Low Power Restricted Boltzmann Machine Using Mixed-Mode Magneto-tunneling Junctions
48	Operations Research	Benson	Christalin	Georgia Institute of Technology	Differentially Private Consensus in Networked Multi-Agent Systems with Optimal Noise Selection
49	Programming Languages	Krzysztof	Drewniak	University of Washington	Synthesizing Data Movement in GPU Kernels Using Abstract Dynamic Programming
50	Robotics	De'Aira	Bryant	Georgia Institute of Technology	Emotional Intelligence in Social Robots designed for Children
51	Security/Privacy	Pablo	Moriano	Indiana University at Bloomington	On Predicting BGP Anomalous Incidents: A Bayesian Approach
52	Security/Privacy	Dieudonne	Mulamba Kadimbadimba	Colorado State University	On Sybil Classification in Online Social Networks Using Only OSN Structural Features.
53	Security/Privacy	Moses	Namara	Clemson University	Privacy Support for Facebook: Empowering Users to Better Control Their Privacy
54	Software Engineering	Miguel	Jimenez	University of Victoria	Software Engineering at Run-time
55	Software Engineering	Luis Felipe	Rivera Vera	University of Victoria	A Vision for Digital Twin
56	Theory	Jesse	Stern	University of Chicago	Volume Only and Multi-Dimensional Reconstruction of Range Query Supporting Database
57	Vision	Chinasa	Okolo	Cornell University	Visual Knowledge Representation of Quorum Sensing

Poster Abstracts

Poster #1 • Brenda Castro
Cornell University

Understanding the Experiences of Teachers of the Visually Impaired with Tactile Materials

Tactile materials, such as tactile graphics and 3D models, are important educational materials for students with visual impairments. Recently, researchers have proposed new methods to improve tactile material creation, but we still have a limited understanding of how tactile materials are created and used today. To address this gap, we conducted semi-structured interviews with 13 teachers of the visually impaired to learn about their experiences teaching with tactile materials. We found that tactile materials were more effective teaching tools than adapted visual materials or verbal descriptions. They also promoted students' independence and cognitive development. However, they were expensive to purchase and time-consuming to create. Furthermore, although 3D models tended to be more effective, teachers often had to use tactile graphics since models were harder to transport and store. Our findings suggest that future research focus on making tactile materials more available, affordable, and portable.

Poster #2 • Sampson Akwafuo
University of North Texas

A Modified Meta-Heuristics Optimization Model for Post-Disaster Emergency Deliveries with Variable Vehicular Capacities

Prompt intervention is essential in reducing casualty figures during sudden epidemic outbreak or emergency situations. This can only be achieved if there is a fit-for-purpose logistics plan in place, incorporating geographical, time and vehicular capacity constraints. In this poster, we present an efficient meta-heuristics algorithm for optimizing logistics and vehicle routing problem (VRP) with variable vehicular capacities. Our algorithm involves two-phased spatial partitioning: vertex clustering and route construction. In the initial phase, locations to be visited are strategically placed in a number of clusters. Membership of each cluster represents locations to be incorporated into a single feasible route. The second phase constructs routes along the clusters, using a priority weighting scheme, from two farthest locations. Geographical, time and capacity constraints, and pruning are heuristically implemented, until an optimized set of solution is obtained. Computational application of our algorithm to test instances in a most-at-risk region clearly indicated its efficiency and accelerated computational execution time.

Poster Abstracts

Poster #3 • Gustavo Aguilar Alas
University of Houston

Dependency-Aware Named Entity Recognition with Relative and Global Attention

Named entity recognition is one of the core tasks in NLP. Although many improvements have been made on this task during the last years, the state-of-the-art systems do not explicitly take into account the recursive nature of language. Instead of only treating the text as a plain sequence of words, we incorporate a linguistically-inspired way to recognize entities based on syntax and tree structures. Our model exploits syntactic relationships among words using a recursive LSTM guided by dependency trees. Then, we enhance these features by applying relative and global attention mechanisms. On the one hand, the relative attention detects the most informative words in the sentence with respect to the word being evaluated. On the other hand, the global attention spots the most relevant words in the sequence. Lastly, we linearly project the weighted vectors into the tagging space so that a Conditional Random Fields classifier predicts the entity labels. Our findings show that the model learns strong correlations between certain words and specific entity types (e.g., verbs such as "eat" and "sleep" are associated with the type PERSON, while prepositions such as "at" and "in" are related to the type LOCATION). Finally, our approach establishes a new state-of-the-art in the SemEval 2010 Task 1 dataset.

Poster #4 • Irish Medina
University of Waterloo

Predicting Short-Term Water Consumption for Multi-Family Residences

Smart water meters have been installed across Abbotsford, British Columbia, Canada, to measure the water consumption of households in the area. Using this water consumption data, we develop machine learning and deep learning models to predict daily water consumption for existing multi-family residences. We also present a new methodology for building machine learning models to predict daily water consumption of new housing developments. This work contains three main contributions: First, we build machine learning models which include a feature engineering and feature selection step to predict daily water consumption for existing multi-family residences in the city of Abbotsford. This is motivated by the recent development direction towards denser living spaces in urban areas. We present the steps of the model building process and obtain models which achieve accurate predictive performance. Second, we present a new methodology for building machine learning models to predict daily water consumption for new multi-family housing developments at the dissemination area level. Currently, the models used in the industry are simple baseline models which can lead to an overestimation of predicted water consumption for new developments, leading to costly and unnecessary investments in infrastructure. Using this new methodology, we obtain a machine learning model which achieves a 32.35% improvement over our best baseline model, which we consider a significant improvement. Third, we investigate the use of deep learning models, such as recurrent neural networks and convolutional neural networks, to predict daily water consumption for multi-family residences. In our case, the main advantage of deep learning models over traditional machine learning techniques is the capability of deep learning models to learn data representations, allowing us to omit the feature engineering and feature selection steps and thereby allowing water utilities to save valuable time and resources. The deep learning models we build achieve comparable performance to traditional machine learning techniques.

Poster Abstracts

Poster #5 • Ademola Okerinde
Kansas State University

Traffic Optimization System using Machine Learning Techniques

"In recent years, motorists on major roads across the world are on a daily basis having a nightmare experience, unfortunately very little attention has been paid to this plight. In this work, we seek to explore this by classifying the experience of motorists as either acceptable or unacceptable. A realistic dataset will be that for which where each motorist is able to review their experience by documenting and reporting their daily experience on each road network. However, this is clearly infeasible, and very few people, if any, will be willing to document and report such enormous amount of data. In this work, we simulated twenty-one intersections in road networks across Poland by measuring the delays caused by red traffic light signal from one intersection to another, and classify users' experience on this basis."

Poster #6 • Brienna Herold
Rochester Institute of Technology

A multi-dimensional method for antigenic characterization of influenza hemagglutinin

Despite ongoing international public health efforts, the influenza virus continues to partially evade vaccine-induced immunity. The influenza virus mutates often, undergoing antigenic drifts or antigenic shifts that make it challenging for current low-throughput methods to rapidly assess individual and population immunity, and to predict who will respond to new vaccines. Influenza vaccines work by inducing long-lasting, type-specific IgG antibodies directed against hemagglutinin (HA). This response is a function of prior IgG-mediated immunity to influenza proteins, as well as the antigenic properties of the influenza strain HA in the vaccine. In this work, we describe an integrated method to predict vaccine responses based on granular antigenic determinant regions, or "epitopes," of the influenza hemagglutinin protein. This method appears promising for both epitope identification in vaccine development and further modeling to predict vaccine induced immune responses.

Poster #7 • Hassan Karim
Howard University

Brain computer interfaces (BCI) - Towards an AI framework for safety, privacy and security considering challenges with algorithmic fairness

If ethnicity or age can be inferred via brain computer interface (BCI) measurements, what AI techniques could be employed to ensure BCI users' safety, privacy and security considering embedded algorithmic fairness challenges. This is an early look at a work in progress. This poster describes some of our early efforts to provide tools which ensure that brain computer interfaces are safe and robust and free of embedded ethnic/racial biases or ageism. Ethnic biases & Ageing are being called out as specific AI features that needed to be investigated since if racial biases or ageism were built into AI models, then BCIs might no longer be safe for certain demographics since using BCIs might subject them to discriminatory practices that produce harm even without the BCI. We present planned experiments to determine the BCI machine learning models needed to extract ethnicity and age biomarkers. We also present use cases where wearable BCIs might need to use ethnicity and age. We also highlight high level steps in our research plan.

Poster Abstracts

Poster #8 • Valeria Velasquez-Zapata
Iowa State University

Next-Generation Screening for Host Interactors of Fungal Effectors in Cereal Immunity

Plant pathogens are among the greatest deterrents to crop production worldwide, resulting in yield losses of 10-20% each year. Many of the aspects of plant defense have been investigated using a traditional gene-by-gene approach. Yet the dynamics of initiation, and the temporal and spatial control of these immune processes are not well understood. Thus, a holistic view of the regulatory programs that render a plant resistant to pathogens is vital to the agricultural economy. To tackle this challenge, we have exploited the biotrophic interaction between the powdery mildew fungus, Blumeria graminis f. sp. hordei (Bgh), and its host, barley (Hordeum vulgare L.). Because pathogens manipulate diverse aspects of plant cell biology to colonize and reproduce within their hosts, effectors represent ideal tools to characterize processes critical to immunity. We use next-generation sequencing to identify interacting partners from high-throughput yeast two-hybrid assays using Bgh effectors as baits, and as preys, a cDNA library from infected barley tissue encompassing key phases of fungal development. The interacting partners are scored using a robust bioinformatics and statistics pipeline, which includes the identification of fusion reads and analysis of count data and these criteria: 1) significant enrichment under selection, 2) in-frame selection, and 3) not enriched with the negative control bait. Scores are used to rank the candidates for further validation tests. The protein-protein interaction data is intersected with already-generated RNA-Seq and expression Quantitative Trait Locus (eQTL) datasets, and a barley protein-interactome to predict the temporal control of host-pathogen interactions at specific stages of infection

---

Poster #9 • Lucca Eloy
University of Colorado, Boulder

Multiparty Regularity in Collaborative Problem Solving

Multidimensional recurrence quantification analysis (MdRQA) quantifies patterns of regularity or repeat states across individuals or signals. We examined these patterns as triads completed a collaborative problem-solving task. We present how this signal of regularity may predict quality and dynamics of triadic collaboration.

Poster #10 • Hirut Kollech
University of Pittsburgh

Biomechanical Response of the Optic Nerve Head Region in Glaucomatous and Non-Glaucomatous Human Samples

"Glaucoma is the second leading cause of blindness in the world. Some of the risk factors for this disease include age, race, gender and intraocular pressure (IOP). More recent studies showed that the biomechanics of the optic nerve head region (ONH) is a critical component for the development of glaucoma.

The purpose of this study is to compare the strain response of glaucomatous and non-glaucomatous human eyes in the lamina cribrosa, which a collagenous structure in the ONH. We used 6 healthy and 7 glaucomatous eyes in the European descent that were above 50 years old. We conducted a pressure inflation experiment using multiphoton imaging and calculated the displacement field using Digital Volume Correlation (DVC). DVC results were used to calculate the Green strain components. Our results show regional variation in both shear and frontal (in-plane) strain between the non-glaucomatous and glaucomatous samples at different pressure levels."

Poster Abstracts

Poster #11 • Hiwot Kassa
University of Michigan

Fitting Demanding Applications to Heterogeneous Systems

"Artificial intelligence (AI) is altering the world in many areas, including business and finance, health care, transportation and many more. The advancement of AI is mainly hastened by fast, powerful computing resources and the availability of massive datasets that can be fed to AI algorithms enabling them to make more accurate decisions. Every day, the data is increasing exponentially and we want to process this data in real time to attain sophisticated AI applications, hence we need computing capacities that grow exponentially. And since we are not getting significant performance improvement from general purpose computing, to handle these demanding applications we need application specific hardware. Application specific hardware have unique computation and datapath that is tailored to the application and algorithm. As AI systems become more complicated it is going to have more algorithms and applications within the system which then requires multiple accelerators working together in one application giving rise to heterogeneous systems.

Heterogeneous systems have two main disadvantages. The first is because emerging applications change frequently, our accelerators are going to be obsolete in a short time which will make the design, testing and verification cost to be very high. The second is when we have more than one accelerator within one system, there should be a programming model that can enable application developers to use different accelerators together within an application.

To solve the first problem our solution is a computation and data access pattern-based accelerator design. Instead of designing accelerators for applications, we propose designing accelerators for common patterns that exist within an application which will increase reusability. For the second problem, our solution is to have a language- and compiler-level framework that can be used to identify different patterns in an application and map different computations in an application to different hardware accelerators. This will create accelerators that can be used by applications with common pattern which is more cost efficient than accelerator per application. Even when an application change we can still map applications to a different accelerator that is close to the changed pattern.

In this research we are identifying the patterns that occur the most in representative AI and machine learning algorithms and we are going to design identification of important patterns and compiler support that can divide code and map it to the accelerator that can run the code efficiently in the system. We are also designing efficient accelerator communication and coordination mechanism because without this no matter how fast the algorithms can run on the accelerators if the communication and data sharing is not efficient we are not going to achieve performance.

Our solutions will help in the advancement of AI algorithms and applications because as applications get diverse and complicated using multiple accelerators in our systems will become inevitable because we are going to have different computation and datapath within an application. Our solutions are going to provide efficient application mapping to heterogeneous system so that application developers then can write programs that will run efficiently on different hardware without the requirement of knowledge about the hardware accelerators and optimization for specific hardware. It is also going to provide accelerator developers information about what kind of accelerators are required and how they should coordinate with one another in resource demanding AI applications to achieve high performance."

Poster Abstracts

Poster #12 • Maria Isabel Mera Collantes
New York University

Don't Trust, Verify: A Verifiable Hardware Accelerator for Matrix Multiplication

In this poster, we propose VeritAcc, a novel approach that enables secure integration of an untrusted third-party matrix multiplication hardware accelerator in a system-on-chip containing a trusted general-purpose processor. Our approach builds upon the theory of interactive proof (IP) protocols to enable run-time verification of each computation executed on the untrusted accelerator and formally guarantees that any incorrect results are detected with high probability. Our hardware implementation contains several novel optimizations to reduce area and performance overhead of VeritAcc. We show that an FPGA prototype of VeritAcc introduces less than 6.25% area overhead compared to a baseline untrusted matrix multiplication accelerator while enabling 11x-69x speed-ups compared to software execution.

Poster #13 • Thierry Tambe
Harvard University

Adaptive Floating-Point Quantization of Deep Neural Networks

Deep Neural Networks (DNNs) have been shown to be resilient to low-bit precision quantization, and as result, enabling significant performance and power gains in DNN hardware accelerators. Recently, the Posit number system has garnered a lot of interest from the deep learning community and is being viewed as a drop-in replacement for the legacy IEEE754 float datatype mainly because Posit can provide a wider dynamic range and more decimal accuracy of the DNN model's parameters compared to Float. However, we observed that Posit efficacy is tightly coupled to a narrow statistical distribution of the model's parameters and becomes less effective than Float in emerging recurrent neural network (RNN) models. We propose an adaptive floating-point-based quantization scheme which yields better prediction accuracy performance compared to Posit, Integer and IEEE-like Float quantization, along with an efficient, reconfigurable, hardware accelerator for sequence-to-sequence RNN networks, leveraging this novel quantization technique.

Poster #14 • Renesha Hendrix
Georgia State University

The Rebirth of Engineering Education

Over the years, universities have researched to implement new ideas to increase the retention rate in their Computer Science departments. However, there is little research on programs to implement in kindergarten through 12th grade (K-12) education to prepare students for higher education computer science courses. To bridge the gap between K-12 and higher education computer science, this study designed and implemented a computer science class in high school. Abstraction was used to create an algorithm from the beneficial factors of education and computer science that was applied to the class. Results from the analysis will be used for further research to improve the course and coursework.

Poster Abstracts

Poster #15 • Kenya Mejia
University of Washington, Seattle

Subject Matter Expertise: Insights on Knowledge Use

This work investigates the role that subject matter expertise plays in the design process. We investigate the design of playgrounds created by individuals with design expertise or engineering expertise in a three-hour lab-based setting.

Poster #16 • Olivia Nche
Clemson University

CodeTracesure - A video game to pinpoint student obstacles and facilitate early detection of students who may need help

CodeTracesure, is a video game which we designed and have used in a summer camp for elementary school African-American children. The game is designed to help children practice CS concepts as they play and to help teachers learn about their difficulties. It also facilitates early detection of students who may need additional help. CodeTracesure covers such concepts as assignment, variables, sequencing, and operators. It is equipped with a database to collect data that could be analyzed for trends and patterns. The current goal is for it to be a supplementary tool that can help the students practice while allowing teachers to collect useful data that can help improve the learning process. An initial study was conducted using this game with about 40 African-American elementary school children. Findings show that the game was useful in motivating the students to practice code tracing and learn CS concepts. The backend end data that was collected on the performance of the students helped to identify potential pitfalls.

Poster #17 • JaCoya Thompson
Northwestern University

Sports Analytics: Learning Data Analytics through the use of Wearable Technologies and Sports

Increasing the numbers underrepresented youth in STEM careers is both an important way to reduce poverty in low income communities, and a contribution to the diversity of thought and experience that drives STEM research. Creating a learning environment, through sports and data analytics, allows students and researchers to meet and explore an area of shared interest. Wearable technologies are promising in different contexts, but continued research is needed to explore their utility in K-12 education. We have the opportunity to expose underrepresented middle school students to wearable technologies and create a fun and engaging hands-on program that draws connections between sports and data analytics. Sports analytics program can provide a rigorous, yet tangible application of math and statistics in which youth perform data gathering and analysis directly linked to their own sports related performance. This allows students who are not traditionally engaged in STEM to be intrinsically motivated to use STEM concepts as a tool for sports training. We will design a learning experience around different types of wearable technologies such as kits featuring conductive thread, LED lights, sensors and other components commonly found in high-tech garments. Students will also work with microcontrollers, which include miniscule circuit boards that can be programmed to direct the various devices attached to them. Activities are designed such that students have access to industry-standard data science tools such as Python statistical packages and R programming tools prominently used by data scientists. The curriculum and associated professional development for the program are designed to encourage connections between in-school and out-of-school time instruction. Providing participants, the opportunity to explore statistical questions that are meaningful to them and allow them to become creators of data science artifacts.

Poster Abstracts

Poster #18 • Amari Lewis
University of California, Irvine

A Preliminary Design for a Blockchain-Based IoT Paratransit System

Public transportation services for users with mobility disabilities has been an important public service in the US for decades. Since the passage in 1990 by the American Disability Act, paratransit has been more visible and important. These services contribute to the mobility of many people with disabilities as well as the elderly. These transportation services are very important but they are notoriously expensive to provide and often inconvenient for users. Recently, many transit agencies and private companies are considering the integration of Transportation Management Company (TMCs – such as Uber, Lyft and Didi Chuxing) with both standard and paratransit services. However, such integration will rely on smart contacts and lately the technical discussion surrounding these contracts has focused on the promise of integrating Intelligent Transportation Systems (ITS), Internet of Things (IoT) and blockchain methods. IoT is a collection of connected devices and sensors that monitor the environment and communicate with each other to accomplish various tasks, typically comprised of low-cost, energy efficient devices. Blockchain networks promote a decentralized environment where each node has a copy of the others on the network. But, through various security measures such as hashing, the identity of the node remains secure. Although ITS has promised to transform transportation systems for decades now, there has not been much progress in paratransit. Thus, we are developing a design for a blockchain-based IoT paratransit system. Within this design, we propose to use smart contracts as the main communication between the users IoT device and the paratransit system. Smart contracts are written in the high-level programming language, Solidity; this is an object-oriented programming language used to deploy the smart contracts amongst the nodes in a non-trusting environment. The ultimate goal of this work is to achieve an ad hoc network environment that involves little human interaction to result in faster route updates, secure ride payments and a better overall experience for paratransit users resulting in a healthy IoT ecosystem.

Poster #19 • Adewale Obadimu
University of Arkansas at Little Rock

Leveraging Social Media Analytics Tools To Understand Information Operations In Modern Communication Platforms

Online Social Networks (OSNs), once regarded as safe havens for sharing information and providing mutual support among groups of people, have become breeding grounds for spreading hoaxes, political propaganda, and radicalizing content. In this charged atmosphere, information – or more frequently its counterpart, disinformation – is emerging as the secret weapon used by adversaries to tip the scales. However, despite the ubiquity of weaponized narratives on various OSNs, there is a lack of systematic research that examines the influence of such operations in modern communication platforms. Most researchers focus on one platform to study disinformation, but dissemination strategies have evolved to a strategic orchestration of different media platforms. This calls for rigorous studies on socio-technical behavioral modeling on OSNs that can explain influence and effectiveness of disinformation campaigns. We start with a comparative analysis of bot activities on Facebook and Twitter. Then we leveraged social media analysis tools to track information actors and narratives as disinformation is disseminated through online social networks including blogs, YouTube, and Twitter. These applications use algorithms that can quickly sift through massive amounts of social media data and identify the right signals that point toward users who are trying to influence behaviors.

Poster Abstracts

Poster #20 • Lucianna Kiffer
Northeastern University

Examining Persistent Hard Forks in Bitcoin and Ethereum

"One of the fundamental properties of major cryptocurrencies to date is the forking upgrade mechanism. If not all players update the code they are running, network partitions occur. In the case of cryptocurrencies these partitions result in competing blockchains, meaning competing versions of the same currency, termed a hard fork. Generally speaking, developers do not risk pushing a new protocol that may result in a persistent fork of the network unless there is some believed off-chain consensus on adopting the new protocol. There are two main exceptions to this rule; the first being controversial forks on a deadline, this was the case with the Ethereum/ Ethereum classic fork, and the second are purposeful forks which have the intention of creating a new coin off the old chain, both Bitcoin and Ethereum have seen many versions of this second kind of fork. We are interested in better understanding what characteristics these forks have and how they may affect the two currencies resulting from them. In order to answer these questions, we plan to use both on-chain data like mining difficulty of blocks and block times, as well as off-chain data like the market values of the currencies.

In the poster I will be presenting our preliminary analysis of the Bitcoin/Bitcoin Cash fork in comparison with our previous analysis of the Ethereum/ Ethereum Classic fork\footnote{Kiffer, Lucianna, Dave Levin, and Alan Mislove. ""Stick a fork in it: Analyzing the Ethereum network partition."" Proceedings of the 16th ACM Workshop on Hot Topics in Networks. ACM, 2017.}. Following a fork, miners in the previous system are now split between the two networks and so there is a stabilization period for the blockchains to catch-up to their new network hash power. We compare the stabilization period by looking at the block rate of each network (seen in the blocks per hour/day in the figure below), and see that where Ethereum classic took about a week to stabilize, Bitcoin Cash took several months. With these forks the two new systems still use the same mining function, i.e. the same hardware, and so miners have the option to choose which currency to mine on. This choice is influenced by the expected return of the miner, we estimate this below using the mining difficulty of each system and the market value of each coin. We note that by the time Ethereum Classic got market value, the system had stabilized and the market value was proportional to the computational power of the network, i.e. equally profitable to Ethereum. On the other hand, Bitcoin Cash had almost immediate market value and had a different scenario where the immediate expected return for mining was sometimes higher for one coin vs. the other. "

Poster #21 • Mahdokht Afravi
University of Texas at El Paso

Long-Term Multiparty Interaction in Virtual Reality

Mahdokht Afravi graduated from the University of Texas at El Paso in 2017 with a B.S. in Computer Science, and a minor in Mathematics. She began her PhD with her advisor Dr. David Novick. in August 2017 at UTEP, where she has continued her undergraduate research in the Interactive Systems Group. Her research experience has included fuzzy logic, computer vision, and human-computer interaction. Her professional goal is to research the use of augmented or virtual reality for fully-immersive environments in long-term multiparty human-ECA interaction.

Poster Abstracts

Poster #22 • Oliver Alonzo
Rochester Institute of Technology

Reading-Assistance Tools for Self-Directed Learning by Deaf and Hard-of-Hearing Computing Workers

Prior research has observed lower levels of English literacy among many Deaf and Hard-of-Hearing (DHH) adults in the U.S. This poses a challenge for DHH workers in computing fields who must regularly teach themselves about evolving technologies. While artificial intelligence techniques can be used to create reading-assistance and text-simplification tools, prior research has not focused on DHH users. In order to adapt these tools to benefit DHH computing workers, we need to understand their needs and preferences, evaluate and prioritize different design parameters, and adapt the tools to fit their literacy profile and to incorporate texts from the computer domain. We also need to understand how these technologies might benefit DHH users and how to measure their efficacy. To do this, we employ qualitative data collection methods, conducting interviews, surveys and a focus group with DHH computing workers and DHH students who plan to work in the field. We will also conduct prototype testing, methodological research on evaluation metrics and NLP research on domain adaptation and syntax-based neural machine translation for text-simplification.

Poster #23 • Larwan Berke
Rochester Institute of Technology

Preferences and Requirements of Deaf and Hard-of-Hearing Users for Captions Generated via Automatic Speech Recognition

"This research project investigates the requirements and preferences of Deaf and Hard of Hearing (DHH) users for captioning technology for both online video programming, including techniques for incorporating Automatic Speech Recognition (ASR) to produce caption text. We are conducting a large-scale study of the diverse DHH community, using focus-groups and survey methodologies, to investigate preferences for new video captioning services, including factors needed to create automatic metrics of video captioning quality. In addition, since ASR is imperfect for the task of automatic caption text production, we are investigating techniques for conveying to users when some words in the automatically generated text output have been recognized at a lower level of confidence (accuracy) by the ASR system. Through experimental evaluations of prototype captioning technologies, we investigate a variety of methods of conveying this confidence information to users, as well as establishing user-interface design parameters for other aspects of caption presentation for users."

Poster Abstracts

Poster #24 • Tya Chuanromanee
University of Notre Dame

Evaluation and Comparison of Usability of Four Mobile Breathing Training Visualizations

Breath training has been shown to reduce stress and anxiety and improve physiological health. Traditionally, teaching breathing techniques involved in-person guidance or using audio recordings. The increasing prevalence of mobile applications has resulted in many different breathing visualization methods. However, limited work has been done in investigating the effectiveness of these visual representations with respect to breathing guidance. We evaluate four common visual representations of the breath used to teach breathing techniques and investigate the role that auditory guidance plays in users' experience of these applications. We develop a mobile interface containing these four representations and use a within-subjects approach. Each participant uses four different types of representations, each with and without audio. Then, we interview and survey participants to gather their thoughts on each representation. We identify potential usability issues with the representations and from the data collected, suggest design criteria for future development of app-supported breath training.

Poster #25 • Steven Díaz Hernández
University of South Florida

Lower Body Rehabilitation System using Wearable Sensors

"According to multiple sources, the number of total knee and hip replacements, limb amputations, and joint related problems has doubled in the last decade. Physical therapists provide patients with post-operative physical therapy to improve their functional outcomes. Unfortunately, most physical therapists either evaluate the rehabilitation process in a subjective manner, or use very expensive and static equipment. Therefore, novel quantitative methods and cost-effective systems using wearable sensors to access and track the progress of patients during rehabilitation process are very important. This research proposes a system that evaluates in a quantitative manner people's gait with the use of wearable sensors. Methods that extract and present gait parameters and symmetry and walking deviation measures from wearable sensors are presented."

Poster #26 • Gabriella Johnson
University of Colorado Boulder

Eye in the Sky: An Audio Description System for Board Games

A majority of board games communicate their information visually, like the location of the game pieces and the information displayed on the cards are acquired through sight. Due to a lack of non-visual feedback, board games pose accessibility challenges to people who are blind or visually impaired as they have to rely on sighted players for information without any way of verifying if the information is accurate, thereby creating a disadvantage. Eye in the Sky is a system that provides an audio description about the state of a board game allowing blind and visually impaired people to play games independently. This system will allow blind or visually impaired people to query for information about the game at any given time, and to strategize their next step without their opponents knowing their next move or plan. This poster will present the current design of the system, which can empower people who are blind or visually impaired to play games independently.

Poster Abstracts

Poster #27 • Lindah Kotut
Virginia Tech

Designing Mobile Technology to Support Hands-Off Interactions

How do we provide hands-on interactivity with an exhibit that one cannot touch? We consider this question when designing for a means of interaction with a garden exhibit housed within a science museum. Our approach leverages mobile phones ubiquity, supporting group interactions and gamification in approaching this conundrum. We then designed a mobile application that leveraged visual recognition in a scavenger hunt game that we deployed and tested in-situ at the museum with children in field trips, teenagers in self-guided explorations, and parent-child dyads in a day-long showcase. We find that compared to non-participants, our design approach notably improved interactions with the garden exhibit and best used the (limited) time available for exploration. We also discuss new insights learned from this atypical exhibit, together with implications on how interaction patterns and group dynamics are changed in the presence of technology.

Poster #28 • Aboubakar Mountapmbeme
University of North Texas

Accessible Blockly: A Block-Based Programming Platform for Users With Visual Impairment and Sighted Users.

Block-based programming languages such as Scratch, Snap and Blockly are commonly used to introduce novices to programming. Because they mostly rely on drag-and-drop actions, they often are not accessible to people with visual impairment, who depend on screen reader software to access information on the screen. Blockly is a popular open-source block-based language designed by Google. Our work focuses on making Blockly accessible to users with visual disabilities while preserving all the features that make this system appealing to sighted users. While Google provides a separate accessible text-based version of Blockly, we believe having separate systems for users with visual impairment and sighted users does not promote collaboration among these groups of people, and it also deprives learners with visual impairment from the beauty of block-based systems that are specifically designed to make learning how to program easy and appealing to novices. Accordingly, our team has extended Blockly to support keyboard navigation, allowing people with visual disabilities perform drag-and-drop actions by using keyboard commands. Screen reader support has also been enabled using WAI-ARIA features and in some cases, blocks are completely redesigned to facilitate access by all users. A self-customizable interface has also been leveraged in our design. Current and future work involve investigating how to leverage audio cues to improve accessibility and convey more information, and conducting large scale usability tests with users with visual impairment and sighted users alike.

Poster Abstracts

Poster #29 • Larry Powell
Texas A&M University

Everyday-Inspired Movies: Movie Recommendation System Based on Personal Social Media Posts

"With the rise of social media platforms, people are providing more of their life and personal information through these systems. With the rise of big data, these often-public social media posts are opening possibilities in many domains such as healthcare. Currently, companies are using such social media information for ad content and recommendations. We propose a system that uses people's personal posts as a means of understanding their real-world situations in life and providing recommendations for movies. Recommended movies will be aligned with each specific viewer's real life and hopefully help them to find the movies more relevant, resulting in a better experience and perhaps even motivate them to reach their goals or supporting them in hardships. The system uses natural language processing to understand people's posts and with machine learning, we recommend the right movie and genre that specific viewers watch. Our work contributes to developing a new personally relevant approach to movie recommendations.

Poster #30 • Isabel Rayas
University of Southern California

Affective State Estimation in Human Interactions with Virtual Agents

In order to have a more effective interaction with humans, robots should be able to recognize human affective states and be able to leverage their actions to influence these states. This research uses Bayesian inference to develop a hidden Markov model (HMM) based on the DAIC-WOZ database from USC's Institute for Creative Technologies, which consists of semi-structured clinical interviews between participants and a virtual agent. The interviews are meant to support the diagnosis of conditions such as depression. With the HMM, we describe the transitions of the human's state through time, and attempt to extract what kinds of behaviors from the virtual agent result in these state changes. This work is foundational to develop a multimodal, real-time, and adaptive model for state estimation and response in human-robot interactions.

Poster #31 • Wendy Roldan
University of Washington, Seattle

The Role of Funds of Knowledge in Online Search and Brokering

Lower-socioeconomic (SES), English Language Learner (ELL), immigrant parents and children work collaboratively to search the internet. Family members rely on each other's language and digital literacy skills in a collaborative information problem solving process known as online search and brokering (OSB). While previous work has identified the ecological factors that impact OSB, research has not yet distilled the specific learning processes behind such family collaborations. From a case study analysis of 3 families, this work explores the funds of knowledge that children and parents rely on as they engage in collaborative learning experiences through OSB. We demonstrate computer supported collaborative learning processes in-the-home are often informal, collaborative, highly social, and highly relevant to solving real-life information access challenges. Our work shows how parents and children draw on their funds of knowledge when they search collaboratively, with and for their family members, to build their collective knowledge of technology and problem solving.

Poster Abstracts

Poster #32 • Jonathan Saddler
University of Nebraska, Lincoln

Understanding Student Comprehension through Eye Gaze

An eye-tracking study conducted in a classroom setting with 17 students in Computer Science is presented. The students were a mix of 12 first-year undergraduates (novices) and five Masters students (non-novices). The tasks involved thirteen C++ programs. Students were asked to answer a comprehension question for each of thirteen programs after reading them. Each program is split into a series of chunks, that logically break down the meaningful parts where eye gazes might hint at cognition about parts of programs useful to solving problems. We analyze these eye gaze patterns across chunks for the stories they tell about how these participants went about searching for clues to solve three types of questions - multiple choice questions about overall program function, fill-in-the-blank questions regarding output, and questions that solicit multi-sentence summaries of large chunks of code. Results and implications of the study are discussed.

Poster #33 • Jean Salac
University of Chicago

An Analysis through an Equity Lens of the Implementation of Computer Science in K-8 Classrooms in a Large, Urban School District

Major metropolitan school districts around the United States are implementing computer science in elementary school classrooms as part of the CS for All (CS4All) initiative. Little is known, however, about the success of such a large-scale rollout, especially in terms of equity. In this study we analyze the performance of 4th grade classrooms completing three modules of an introductory computational thinking curriculum, looking at not only overall results but also the variance in performance between high-, mid-, and low-performing schools (as identified by their school report cards). We find that all classrooms are benefiting from the computational thinking (CT) curriculum, making great strides in providing equitable access to CT education. However, statistically-significant differences in performance are present, especially between the high- and low-performing schools, showing that there is still room for improvement in developing strategies and curricula for struggling learners.

Poster #34 • Jomara Sandbulte
Pennsylvania State University

Exploring Photo Sharing to Engage Intergenerational Families on Health

Sharing photos is a common practice for family members to inform and connect with each other. Sharing photos inspires recollection, reminiscence, and feelings of connectedness. Could it also be an effective channel for sharing health information and supporting healthy activity? In this study, we examine how photo-sharing practices can help family members, specifically elderly parents and adult children, share health information. We present preliminary results from an interview study investigating photo-sharing benefits for non-collocated elderly parents and their adult children to encourage healthy lifestyles within the family. We discuss photo sharing as a viable strategy to engage family members in sharing health information.

Poster Abstracts

Poster #35 • Ather Sharif
University of Washington

OptiGain: Automatically Discovering the Optimal Mouse Gain Setting for People With Limited Hand Function through Everyday Computer Use

The mouse cursor control-display (C-D) gain determines how much movement of the mouse cursor results from corresponding physical movement in the world. Colloquially, C-D gain refers to the "sensitivity" of the mouse cursor. For most computer users, this sensitivity is an afterthought, but for users with motor impairments causing limited hand function, the C-D gain can make a big difference in their ability to work efficiently at a computer. At the same time, most users are not aware of the C-D gain setting and, even if they are, might not pick the optimal value. To address these concerns, we are developing OptiGain, a software tool that runs in the background to analyze everyday pointing movements and adjust the C-D gain setting to optimize throughput, a measure of pointing efficiency. As part of this work, we also conducted a study that verified that optimal C-D gain settings do indeed exist for users. The results of this work and our OptiGain tool will help people with limited hand function to perform pointing tasks with increased efficiency and comfort.

Poster #36 • Mark Yousef
University of Minnesota

Measuring Psychological Effects of Peer-to-Peer Reputation Systems Involving In-Person Exchanges

"Reputation systems such as those used by peer-to-peer services are important to help companies filter out problematic consumers and producers. However, consumers may not always be aware of the presence of these systems in the services they use and if so may not always agree with them. Furthermore, it has long been predicted that scores from a variety of services may soon be aggregated into reputation banks. Seemingly the new credit scores for the digital economy, these personal rating systems have unexplored consequences on consumer psyche. Using a case study of Uber passengers, this study examines stress and anxiety levels associated with personal rating scores, as well as their overall acceptance.

Poster #37 • Mayra Samaniego
University of Saskatchewan

Detecting Suspicious Transactions in IoT Blockchains for Smart Living Spaces

"The idea of connecting physical things and cyber components to enable new and richer interactions is a key component in any smart space concept. One of the central challenges in these new smart spaces is the access control of data, services and things. In recent years, Distributed Ledger technology (DLT) like Blockchain Technology (BCT), emerged as the most promising solution for decentralized access management. Using capability-based access control, access to data/services/things is achieved by transferring tokens between the accounts of a distributed ledger. Managing how the access tokens are transferred is, of course, a major challenge. Within the IoT space, smart contracts are at the center of most of the proposals for DLT/BCT networks targeting access control. The main problem in using smart contracts as a means for checking if and what access token can be transferred from one account to another is their immutability and accessibility. Smart contracts and chain code are by design meant to be immutable since they represent a binding contract between parties. In addition, they need to be accessible since they are to be executed on many nodes. This allows an attacker to study them and design the attack in a manner that passes the rules of the smart contract/chain code. This paper focuses on the use of metadata as a more effective means to prevent attackers from gaining access to data/services/things in a smart living space."

Poster Abstracts

Poster #38 • Abiola Arise
Kansas State University

Identifying Features Important for Predicting Information Diffusion in Twitter Events

"Online social networks (OSN) have become increasingly crucial for the dissemination of important information, as well as disaster crises reporting. While the dissemination of accurate information may protect the general public and potentially save lives, the risk of spreading false or inaccurate information may be detrimental to public health and safety in those contexts. In a time when the spread of misinformation is overly populous, it is important to study the creation, spread, as well as opinion formation in OSN. To effectively establish this phenomenon, it is essential that we identify the key features that contribute to the re-post and eventually, the spread of information in OSN.

We hypothesize that the features that contribute to information diffusion in online social networks are significantly influenced by the type of event, based on this, we classify Twitter events into 2 categories: (1) elucidative and (2) trending.

In this study, we propose a mathematical approach to understanding and modeling the features that directly impact information diffusion in informational and trending events on Twitter. First, we identify key user and message attributes that contribute to information diffusion by carrying out node-to-node influence analysis. This process allows us to identify influential nodes that are crucial to maximizing information spread or minimize misinformation. Second, we adopt Bayesian logistic regression to predict the diffusibility of a message given the associated network, interaction, context, and temporal properties. Finally, run Random Forest to select optimal features necessary to achieve maximum predictability."

Poster #39 • Kiante Brantley
University of Maryland

Non-Monotonic Sequential Text Generation

"Standard sequential generation methods assume a pre-specified generation order, such as text generation methods which generate words from left to right.

In this work, we propose a framework for training models of text generation that operate in non-monotonic orders; the model directly learns good orders, without any additional annotation.

Our framework operates by generating a word at an arbitrary position, and then recursively generating words to its left and then words to its right, yielding a binary tree. Learning is framed as imitation learning, including a coaching method which moves from imitating an oracle to reinforcing the policy's own preferences.

Experimental results demonstrate that using the proposed method, it is possible to learn policies which generate text without pre-specifying a generation order, while achieving competitive performance with conventional left-to-right generation."

Poster Abstracts

Poster #40 • Adedolapo Okanlawon
Kansas State University

Cloudbursting: Stay-or-Go?

"Users are known to overestimate or underestimate High-Performance Computing (HPC) job resources (CPU, memory, time, cores) when they submit jobs and this can result in wastage of resources, and consequently, jobs being killed. There is a need to investigate how job submission details available at runtime can be used to determine whether a submitted job will complete or not.

We focus more on jobs whose resources are underestimated and are likely to fail. We hypothesize that prior probabilities from previous job submissions can be used to generate better probabilities based on user history, and then recommend to users to either stay on the local HPC platform or go, that is, move the job to cloud services like Amazon Drive.

We collected log files from an HPC cluster and performed logistic regression, using Gaussian Naive Bayes for estimation of posteriors, on the data to see how job submission information can be used to predict if jobs failed or not. We then classify the tasks into two: (A) A static task where the user only submits the job once (B) A dynamic task in which the user resubmits jobs up to an allowable number of resubmissions (10 in this study). We then derive a cost function to predict the 'job failed' probabilities and hypothesize that the latter approach would perform at least as well as the former.

Poster #41 • Deontae Pharr
Georgia State University

Recurrent Convolutional Neural Networks for Diagnosing Mental Disorders

Diagnosing mental disorders is a considerably complex task for behavioral health professionals due to the many factors that may complicate the process. The current methods that are currently undertaken requires a significant amount of time that doesn't guarantee reliable results since mental disorders often overlap with one to many others, as well as the fact that people exhibit individual behaviors with the symptoms of their disorder(s). To alleviate these issues, utilizing methods that produces objective results from sophisticated technology is essential. Thanks to the functional Magnetic Resonance Imaging (fMRI), a sophisticated Neuroimaging solution, the structure and function of the brain can be detailed and analyzed by professionals for answers to their many medical queries. With this as my premise, I utilize data from an fMRI machine to evaluate the brain of a subject with a mental disorder and attempt to classify it with Deep Learning models, specifically Recurrent Convolutional Neural Networks.

Poster Abstracts

Poster #42 • Angela Stewart
University of Colorado, Boulder

I Say, You Say, We Say: Using Language to Model Shared Knowledge Construction during Collaborative Problem Solving

Collaborative problem solving (CPS) is a crucial 21st century skill, however, cur-rent educational technologies fall short of supporting effective CPS processes, especially for virtual interactions. In order to develop intelligent systems that enhance CPS processes and outcomes, we investigate automated detection of a critical CPS process – construction of shared knowledge. Our data consists of 32 triads who engaged in collaboratively solving a challenging computer programming task for 20 minutes via videoconferencing software. We used automatic speech recognition to generate transcripts of 11,162 utterances, which human experts coded for evidence of shared knowledge construction. We aimed to automate the expert-codes in a team-independent fashion from the words themselves (bag of n-grams) or word categories using standard dictionaries (LIWC categories). Despite imperfect automatic speech recognition, the n-gram and LIWC-category models achieved AUROC (area under the receiver operating characteristic curve) scores of .85 and .82, which reflects a 70% and 64% improvement over chance (AUROC of 0.5). Further, the model-derived scores predicted individual learning similar to expert codes (neither predicted subjective perceptions). We discuss embedding our models in collaborative interfaces for assessment and dynamic intervention with an eye for improving CPS outcomes.

Poster #43 • Angelina McMillan-Major
University of Washington

Automating Linguistic Annotations in IGT for Low-Resource Languages

The majority of the speaker communities of the world's languages are generally unable to benefit from the machine learning techniques tailored to extracting and applying linguistic knowledge because they do not have the necessary quantity of digitized data required by these techniques. Interlinear Glossed Text (IGT) is often the one of the few ways in which a low-resource language appears online, thanks to the work of field linguists and their native-speaker collaborators. IGT is a format for presenting a sample of a language, its linguistic analysis, and its translation in the target audience's language. Using the ODIN database of enriched IGT, I combine rule-based linguistic knowledge and statistical machine learning to augment linguistics annotations for low-resource languages. This generative system is trained with just under 1,000 training instances for each language independently, where the input is the segmented source language phrase and its (generally English) translation. The output is the predicted linguistic annotation for the source language phrase. The accuracy of the system varies across languages based on linguistic typology as well as annotation consistency.

Poster Abstracts

Poster #44 • Lena Reed
University of California, Santa Cruz

Can Neural Generators for Dialogue Learn Sentence Planning and Discourse Structuring?

Responses in task-oriented dialogue systems often realize multiple propositions whose ultimate form depends on the use of sentence planning and discourse structuring operations. For example, a recommendation may consist of an explicitly evaluative utterance e.g. Chanpen Thai is the best option, along with content related by the justification discourse relation, e.g. It has great food and service, that combines multiple propositions into a single phrase. While neural generation methods integrate sentence planning and surface realization in one end-to-end learning framework, previous work has not shown that neural generators can: (1) perform common sentence planning and discourse structuring operations; (2) make decisions as to whether to realize content in a single sentence or over multiple sentences; (3) generalize sentence planning and discourse relation operations beyond what was seen in training. We systematically create large training corpora that exhibit particular sentence planning operations and then test neural models to see what they learn. We compare models without explicit latent variables for sentence planning with ones that provide explicit supervision during training. We show that only the models with additional supervision can reproduce sentence planning and discourse operations and generalize to situations unseen in training.

Poster #45 • Andrea Villanes
North Carolina State University

Epidemiological Disease Surveillance Using Public Media Text Mining

Despite the improvement in health conditions across the world, communicable diseases remain among the leading mortality causes in many countries. Combating communicable diseases depends on surveillance, preventive measures, outbreak investigation, and the establishment of control mechanisms. However, delays in obtaining country-level data of confirmed communicable disease cases, such as dengue fever, are prompting new efforts for short- to medium-term data. In this work, we propose the creation of a surveillance tool for communicable diseases, with a focus on dengue fever, by analyzing data on public media. Our research offers the following novel contributions to text analytics, sentiment analysis, epidemiology, and visualization areas: (1) an alternative method for near real-time estimation of disease outbreak, spread, and response based on text analytics of public media sources like newspapers and social media; (2) identification of topics extracted from epidemiological news articles using text mining cluster analysis and topic analysis, which has not been used before in public health surveillance systems; (3) comparison of existing text mining classification techniques to accurately predict news article topics; (4) creation of a communicable disease sentiment dictionary by extending an existing dictionary with epidemiological terms and their associated sentiments, this sentiment dictionary can be used to estimate sentiment in the area of public health; (5) creation of a streamgraph inspired technique to display evolution of topics over time, incorporating known trends to allow for comparison; (6) integration of our cluster, categorization, sentiment analysis, and visualization techniques into an interactive web-based tool that allows domain experts to monitor dengue fever. This tool can be used as the basis for monitoring other communicable diseases in the future.

Poster Abstracts

Poster #46 • Ranysha Ware
Carnegie Mellon University

Battle for Bandwidth: Fairness and Heterogenous Congestion Control

"Congestion control algorithms (CCAs) at every host on the Internet are responsible for determining how fast to transmit at. One important responsibility of CCAs is ensuring that congested links are shared according to some definition of fairness—without a few senders dominating the link capacity and crowding others out. In recent years, the original CCA, TCP Reno, has been replaced by a suite of very different approaches including TCP Cubic in Linux, Google's BBR, and Akamai's FastTCP. This change in CCAs is significant because, with so many new CCAs, the networking community does not know how these heterogeneous algorithms fairly (or unfairly) share the network.

In this work, we propose a new testing methodology for measuring the impact of a production Internet service's CCA on the performance of competing traffic. We conduct a census of the CCAs deployed in the Internet today. We find widespread deployment of BBR and a handful of unknown CCAs. We then apply our new testing methodology and find that services using BBR can be extremely unfair to TCP Cubic. We also see unfair outcomes between TCP Cubic and a service using an unknown CCA. Our work shows that heterogeneity is the standard (rather than the exception) on the Internet today, and defines a better methodology to decide if new CCAs will be fair to other senders on the Internet; our goal is that by sharing these results we will contribute to a more equitable Internet for all users."

Poster #47 • Shamma Nasrin
University of Illinois at Chicago

Low Power Restricted Boltzmann Machine Using Mixed-Mode Magneto-tunneling Junctions

This work discusses the mixed-mode magneto tunneling junction (m-MTJ)-based Restricted Boltzmann Machine (RBM). RBMs are unsupervised learning model, suitable to recognize, cluster and generate images, video sequences, and motion-capture data. We can also build a deep belief network by stacking RBMs. m-MTJ is a two terminal device actuated by the simultaneous actions of voltage-controlled magnetic anisotropy and voltage-controlled spin-transfer torque, where the switching of the free-layer is probabilistic and can be controlled by the two. Using m-MTJ based activation functions, we present a novel low area/power RBM neuron. Each node of RBM has a probability to be in state 1. A low area/power RBM is desired, for example, to extract low-dimensional features in a sensor node itself, thereby, averting the transmission of voluminous raw data from a sensor node to cloud. Unlike a typical neural network, the activation functions in RBM are stochastic and produce an output `0' or `1' with a probability determined by the network weights and the applied pattern. However, a physical implementation of stochastic activation function in RBM is challenging. E.g., previously, an ASIC implementation of the activation functions consumes 137.9 μW power. This work discusses the potential of mixed-mode magneto-tunneling junctions (m-MTJ) to simplify the activation functions in RBM. In this work, we also present a memristor crossbar architecture to implement the weights and MAC operation of the network. We also discuss an online training of RBM to negate the impact of process variability on its functional accuracy. A device level simulation based on physics relation is used to validate the sigmoidal probabilistic output of m-MTJ with an applied voltage across it. Characteristic of the memristor crossbar and m-MTJ are modeled in HSPICE to investigate the performance and power tradeoffs of the network. In application-level simulations, the RBM is implemented in MATLAB for digit recognition. For handwritten character recognition on MNIST, the design achieves 96.6% accuracy under expected variability in design components.

Poster Abstracts

Poster #48 • Benson Christalin
Georgia Institute of Technology

Differentially Private Consensus in Networked Multi-Agent Systems with Optimal Noise Selection

This project demonstrates the initial phase of developing an algorithm that preserves differential privacy for a consensus multi-agent networked systems problem with an emphasis on performance guarantees. The objective is for agents of a multi-agent control system to flock together by using state information from their neighbors while keeping these values differentially private against an adversary that has access to the messages. The differentially private algorithm corrupts messages with Laplacian noise and is guaranteed to achieve average consensus in expectation. We examine how to optimally select the noise parameters to minimize the variance of the network convergence point and time constant for convergence for the desired number of agents and level of privacy.

Poster #49 • Krzysztof Drewniak
University of Washington

Synthesizing Data Movement in GPU Kernels Using Abstract Dynamic Programming

Many GPU kernels, such as the inner loops of a convolution, obtain their high performance from patterns of inter-thread data movement that are difficult for humans to write. These kernels can be generated by synthesizing swizzles, which are expressions that determine how threads will communicate with each other, that are left unspecified in a program sketch. This delegates the desk of determining the communication patterns from a human to an algorithm, greatly simplifying kernel programming. We present a novel synthesis algorithm that prunes an enumerative search using dynamic programming on abstractions of candidate solutions. We have successfully applied this method to kernels that permute their inputs and have obtained significant speedups compared to previous methods based on satisfiability modulo theories (SMT) solvers, and are exploring further applications.

Poster Abstracts

Poster #50 • De'Aira Bryant
Georgia Institute of Technology

Emotional Intelligence in Social Robots designed for Children

Social robots are robots designed to interact and communicate directly with humans, following traditional social norms. However, many of these current robots operate in discrete settings with predefined expectations for specific social interactions. In order for these machines to operate in the real world, they must be capable of understanding the multiple factors that contribute to human-human interaction. One such factor is emotional intelligence. Emotional intelligence allows one to consider the emotional state of another in order to motivate, plan, and achieve one's desires. One common method of analyzing the emotional state of an individual involves analyzing the emotion displayed on their face. Several artificial intelligence (AI) systems have been developed to conduct this task. These systems are often classifiers trained using a variety of machine learning techniques which require large amounts of training data. As such, they are susceptible to biases that may appear during performance analyses due to disproportions existing in training datasets. Children, in particular, are often less represented in the primary datasets of annotated faces used for training such emotion classification systems. This work seeks to first analyze the extent of these performance differences in commercial systems, then to develop new computational techniques that work to improve emotion classification performance in children, and to finally present a social robot which utilizes an improved emotional AI to interact with children in various scenarios where emotional intelligence is key to successful human-robot interaction.

Poster #51 • Pablo Moriano
Indiana University at Bloomington

On Predicting BGP Anomalous Incidents: A Bayesian Approach

Despite multiple efforts to secure the Internet control plane, Border Gateway Protocol (BGP) anomalous incidents have been increasing in both frequency and impact over the past several years. These anomalies are events in which Internet traffic is accidentally or maliciously routed incorrectly. Here, we examine a popular public dataset of BGP anomalies to develop Bayesian Generalized Linear Models (BGLMs) that capture the frequency and impact of BGP anomalies. We find that the daily frequency can be modeled by a lognormal distribution while their impact is better captured by a discrete Laplace distribution. Knowing these distributions can provide insights into the generative mechanisms of these anomalies and inform future predictions.

Poster Abstracts

Poster #52 • Dieudonne Mulamba Kadimbadimba
Colorado State University

On Sybil Classification in Online Social Networks Using Only OSN Structural Features.

Sybil attack is a problem that seriously affects Online Social Networks (OSNs). These attacks are made possible by the openness of OSN platforms that allows an attacker to create multiple fake accounts (called Sybils), which are then used to compromise the underlining trust pinnings of the OSN. Early Sybil account detection mechanisms involved classification of users into benign and malicious based on various attributes collected from the user profiles. One challenge affecting these classification methods is that user attributes can often be incomplete or inaccurate. In addition, these classification methods can be evaded by sophisticated attackers. More importantly, user profiles can often reveal sensitive user information that can potentially be misused causing privacy violation. In this work, we propose a Sybil detection method that is based on the classification of users into malicious and benign based on the inherent topology or structure of the underlining OSN graph. We propose a new set of structural features for a graph. Using this new feature set, we perform several experiments on both synthetic as well as real-world OSN data. Our results show that the proposed detection method is very effective in correctly classifying Sybil accounts without running the risk of being evaded by a sophisticated attacker and without compromising privacy of users.

Poster #53 • Moses Namara
Clemson University

Privacy Support for Facebook: Empowering Users to Better Control Their Privacy

Prior research has shown that Facebook users' engagement and use of privacy features greatly differs. Users' find it laborious to translate their desired privacy preferences into particular interface actions. In this study, we probe how User-Tailored Privacy (UTP) can be utilized to tailor Facebook's privacy features to user's personal preferences. Using a "think-aloud" semi-structured interview approach (N=18), we assess how three adaptation methods: Automation, Highlight and Suggestion can be used to suitably tailor Facebook's interface to these personal preferences. Our findings provide awareness about the viability of UTP on Facebook and other social network platforms. In particular, we find that the optimal adaptation method depends on familiar users are with the privacy feature and how they use them paired with their judgement of the awkwardness and irreversibility of the tailored privacy functionality.

Poster #54 • Miguel Jimenez
University of Victoria

Software Engineering at Run-time

New innovative development approaches are required to deal with the increasingly shorter and more frequent software evolution cycles that deliver added-value continually. Despite effective approaches, such as DevOps and continuous software engineering (SE) practices, run-time issues that cannot always be anticipated at design-time still hamper the value and service quality delivered to customers. Researchers in the SEAMS community have addressed many of these issues from autonomic computing and models at run-time considering both sides, development and operations. Nevertheless, we posit that software development requires more innovation in adaptive methods that effect not only the run-time phase but also other phases of the development life cycle. We envision to inject adaptive methods and techniques into the continuous engineering processes to realize what we call SE at run-time. This poster illustrates our vision for continuous SE using the concept of adaptive continuous experimentation and evolution.

Poster Abstracts

Poster #55 • Luis Felipe Rivera Vera
University of Victoria

A Vision for Digital Twin

The Digital Twin concept is enabling new research opportunities in the domain of software engineering (SE) methods, processes, and tools. It offers a novel approach for virtually representing both the structural elements and dynamics of an asset throughout its entire lifecycle. This notion implies a causal connection between each asset and its virtual representation, hence Models at Runtime (MARTs) become fundamental. In this poster, I analyze the future role of MARTs for the design and operation of digital twins. I emphasize in the suitability of MARTs to virtually represent highly dynamic entities for which modeling and comprehending their behavioral aspects becomes essential when understanding evolution over time. As an application scenario, I examine the use of digital twins to not only track patients' health, a highly dynamic entity, but also to evaluate the application and evolution of medical treatments in the healthcare realm.

Poster #56 • Jesse Stern
University of Chicago

Volume Only and Multi-Dimensional Reconstruction of Range Query Supporting Database

Recent work has shown that secure-outsourced database systems that leak range query communication volume can result in full reconstruction of the database. Attacks proposed in KKNO16 and GLMP18 successfully perform database recovery in a passive attack model using leakage functions present in all currently existing order preserving encryption schemes that support range queries. These two methods, although successful, utilize expensive polynomial factoring algorithms and clique-finding algorithms, respectively. Moreover, they are not able to succeed if the queries are restricted to "shorter" lengths. The polynomial factoring method in KKNO16 requires that the adversary observes the volumes associated with each query; this includes queries that range the whole or almost the whole database. Similarly, the graph-algorithmic method presented in GLMP18 requires the collection of all volumes. To date, full reconstruction using only short queries, as more commonly seen in practice, is still an open question. In this paper, we address a method that is able to reconstruct the database using a polynomial time algorithm. We also explore why and in which cases these suggested attacks fail. We also investigate the problem of reconstructing both dense and sparse k-dimensional databases that support k-dimensional range queries.

Poster Abstracts

Poster #57 • Chinasa Okolo
Cornell University

Visual Knowledge Representation of Quorum Sensing

"Quorum sensing is a cell-cell communication process that bacteria use to transition between individual and social lifestyles. It is also a system of stimuli and response correlated to population density. In this system, bacteria can share information about cell density and adjust gene expression accordingly. Vibrio cholerae and Vibrio harveyi are two of many organisms that exhibit quorum sensing behavior.

Current graphical model representations for Vibrio (such as those shown in academic papers) are unable to represent complexities of quorum sensing and are often inconsistent in labeling. These graphical models were the motivation for this project and represent the knowledge of information retrieved from papers about quorum sensing and molecular communication systems.

These incomplete representations also spurred the need to develop better representations of quorum sensing pathways. To do this, it was necessary to turn the information gained from the relationships between proteins and receptors in each respective quorum sensing pathway into a knowledge base encoding the negative and positive interactions that occur. From this, a software implementation of the knowledge base can be integrated within the current framework of the Genetic Engineering of Cells (GEC) project. Altogether, a total of three different representations of the quorum sensing models in Vibrio harveyi were created, encapsulating the current understanding of this respective system.