CCSCNE 2020

Student Poster Abstracts

The Consortium for Computing Sciences in Colleges

The Twenty-Fifth Annual
Consortium for Computing Sciences in Colleges
Northeastern Conference

Scheduled to be held from April 17 - April 18, 2020 at

Ramapo College of New Jersey
Mahwah, New Jersey

Conference postponed until 2021

Poster 5001

Trusted Digital Identity Management using Blockchain

Zacharie Happel, Kaylin Moss

Computer Science and Mathematics
Marist College, Poughkeepsie, NY

Faculty Advisor

Brian Gormanly
Marist College, Poughkeepsie, NY

In the modern digital age, we are faced with a variety of security concerns related to identity theft, impersonation, and forgery of electronic identification information. In this paper, we discuss a proposal to replace physical ID cards with a secure digital alternative that could be managed with a smartphone. The approach is based on IBM's decentralized framework known as Trusted Identity, and in particular a software application called Verify Credentials, which acts as a digital wallet. Personal identifying information is stored on a Blockchain database, which is inherently tamper resistant. We describe a pilot program between IBM and Marist College, the exclusive academic partner for Trusted Identity, to replace college student ID cards with digital identities. Currently, students are issued a physical ID card with Bluetooth and magnetic strip capabilities that includes the student's name, photo, and a unique identification number; this information is authenticated from a centralized database service. Collaborating with stakeholders at Marist, we developed three initial use cases for digital identities: access to campus buildings, password-less authentication into the college's online student portals, and online credential verification for student memberships at a nearby gym. We provide an overview of the new identity management system and results of testing in the initial pilot phase. The project is scheduled to expand to a test group of up to 20 students in 2020, using the Verify Credentials mobile application in a closed environment. We are developing new credential schemes and adapting the architecture necessary to integrate digital identifies with legacy campus security systems. The project goal is to strengthen campus security while decreasing costs and enhancing the student user experience.

Poster 5002

The Bloating of Texting Applications: *WuChat,* One Application to Unite Us All

Jared Zimmerman, Dakota Nickel, Ben Harvilla
Computer Science Department
Widener University, Chester, PA

Faculty Advisor
Suk-Chung Yoon
Widener University, Chester, PA

With the variety of texting methods students have, *Widener University*, like many other universities around the US, has become over-saturated with these applications with no consistent method for students to contact one another, effectively alienating each other. To resolve this over-saturation, we aim to use our application. This project, titled WuChat, is a texting service application that will be available for Android devices, consisting of Java code for the UI housed by *Android Studio*, connection with a *Firebase* housed database for user name organization and storage, as well as password and unique user id encryption. The aim is to create a texting application that will mimic similar apps, such as *GroupMe*, *Kik*, *SMS*, and the texting features of *Snapchat*. We hope that it will become a campus-wide and campus-exclusive texting application aimed toward students, allowing them to have a University monitored method of being in contact with one another, as opposed to using their cluttered school-emails or needing to collaborate on using a specific app to communicate. The current plan is to complete the application completely by June of 2020, and to get endorsement by Widener University.

Poster 5003

Designing a Logic Problem Solver Using a Natural Language System

Aaron Cogbill
Computer Science Department
Widener University, Chester, PA

Faculty Advisor
Suk-Chung Yoon
Widener University, Chester, PA

We present our experience developing a natural language reasoning system at Widener University. The zebra puzzle is a logic problem in which the solver is asked to match pet owners with pets and houses while satisfying arbitrary constraints. We outline a natural language interpreter that generates a solution from an English statement of the zebra puzzle, bypassing the need to restate the problem in a domain-specific language. Our system consists of a semantic parser, a model generator based on Prolog's default proof search, and an interactive shell. We discuss the handling of quantifiers, including the unique quantifier "the," in Prolog, which prove essential for solving the puzzle. We also identify deficiencies in our parsing and model-building system that limit our ability to describe the puzzle to the system and find workarounds. Potential improvements include support for pronouns and other words that reference preexisting logical constants, representation of semantic ambiguity via under-specification, and a more sophisticated model generator/automated reasoning system. While our system is specifically tested on the zebra puzzle, the same concept could be extended to similar types of constraint problems by means of a larger lexicon and grammar.

Poster 5004

Exploring Computer Security Labs and Virtual Learning Environments

Cole Bryner
Computer Science
Widener University, Chester, PA

Faculty Advisor
Yana Kortsarts
Widener University, Chester, PA

Cybersecurity is one of the fastest-growing career fields today and cybersecurity professionals are in high demand due to continuing growth in cybercrime. National Initiative for Cybersecurity Careers and Studies (NICCS), the official website of the Department of Homeland Security (DHS), defines cybersecurity as "the activity or process, ability or capability, or state whereby information and communications systems and the information contained therein are protected from and/or defended against damage, unauthorized use or modification, or exploitation." The current project was focusing on exploration of various computer security virtual learning environments and hands-on labs that potentially could be integrated into CSCI 392 Computer Security course curriculum. Computer Security course is an upper level computer science technical elective and one of the required courses for Computer Forensics minor. For the purposes of this project three computer security environments were examined: Jones & Bartlett Virtual Security Cloud Labs, which were used during first run of the course during Spring 19; NSF-funded SEED hands-on labs for security education; and IBM security labs. The analysis of these computer security educational tools was performed taking in account the cost, the comprehensive coverage of the computer security topics and the level of difficulty to use these environments.

Poster 5005

Detecting Network Anomalies using Machine Learning

Matthew Grubelic, Chris Saliby, Anthony Saldana , Luke Turbert, Andrew Rittenhouse
ECECS
University of New Haven, West Haven, CT

Faculty Advisor
Liberty Page, Valid Behzadan
University of New Haven, West Haven, CT

The risk of a crippling cyber-attack on computer systems is increasing rapidly each day. Current software and techniques used to defend against these malicious attacks are showing their limitations and are being completely overwhelmed in some cases. As a result, a more modern and forward-thinking solution is becoming increasingly necessary. The team is implementing one such software solution using a sequence to sequence neural network in Python3 with the Pytorch library to observe malicious events and predict when the next attack might happen. In this project, a deep learning sequence to sequence model, modeled after the behavior of predictive typing technologies, was implemented on the CICIDS 2017 dataset to detect malicious traffic and determine the probability of another attack in the future. The model functions by observing sequences of network packets, using them to predict upcoming sequences of packets, and comparing the actual observed data to the prediction. Since the amount of notable research and experimentation done in this area so far is lacking, the results yielded by this network traffic anomaly detection approach further demonstrate that the use of a sequence to sequence model is a viable, though still emerging, solution for modern intrusion detection systems.

Poster 5006

Bayesian Phylogenetic Inference of Stochastic Block Model on Random Graphs

Yanqiu Guo

Department of Mathematics and Computer Science
Queensborough Community College, Queens, NY

Faculty Advisor

Wenjian Liu
Queensborough Community College, Queens, NY

Determining the reconstruction threshold of a broadcast models on random graphs, as an interdisciplinary subject, has attracted more and more attention from probabilists, data scientist, statistical physicists, biologists, etc. This project analyzes a classification problem on a deep network, by considering a broadcasting process on an infinite communication tree, where information is transmitted from the root of the tree to all the vertices with certain probability of error. The tree reconstruction problem is to determine whether symbols at the *n*-th level of the tree contain non-vanishing information about the root, as n goes to infinity. Its connection to the clustering problem in the setting of the stochastic block model, which has wide applications in machine learning and data mining, has been well established. For the stochastic block model (SBM), an "information-theoretically-solvable-but-computationally-hard" region, or say "hybrid-hard phase", appears whenever the reconstruction bound is not tight of the corresponding reconstruction on the tree problem.

Inspired by the recently proposed "*q_1+q_2*" SBM, the major task of this project is to extend the classical works on the Ising model and the Potts model, by studying a general symmetric model which incorporates the characteristics of both Ising and Potts through different in-community and out-community transition probabilities, and rigorously establishing the exact conditions for the non-tightness of the reconstruction bound. Furthermore, the mentee will concentrate on the phase transition of the reconstructability of this SBM by establishing the distributional recursion and moment recursion by analyzing the recursive relation between nth and (*n+1*)th generation's structure of the tree. The last work of the mentee is to apply numerical analysis and MATLAB to compute rigorously a combination of basic arithmetic operations and simulate the signal propagation on random graphs.

This is an undergraduate research project funded by the Center for Undergraduate Research in Mathematics (NSF grant).

Poster 5007

The Application of Convolutional Neural Networks – Donkey Car

Markeilia Cassis

Computer Science
New Jersey City University, Jersey City, NJ

Faculty Advisor

T. Liu
New Jersey City University, Jersey City, NJ

The topic of inquiry for this experiment was convolutional neural networks (CNN) and their application to daily life activities such as virtual self-driving. The research's purpose was to prove if reinforced learning (RL) is an efficient way to utilize artificial intelligence (AI). In specific, is it possible to simplify neural networks for the broader use in companies and the daily life of consumers with minimal source code alteration? Donkey Car is our newly assembled vehicle model from the bought parts. Part of the learning process includes the installation of Raspberry Pi software from a laptop. By using Linux and the program's terminal, the Donkey Car's Raspberry Pi was accessed; hence, the car's autonomous functions (including steering) were available to the programmers. After which the car was connected to the network to calibrate steering and throttle through a repeated process of manual test drives. Following this, the data acquisition and the training of the neural networks are done. The embedded camera is used to capture the images of the training tracks. This machine learning training process can be repeated so the accuracy of performance can be improved. Finally, the self-driving accuracy test scores how well the Donkey Car maneuvered on the given track. The results proved that the Donkey Car 2.0's source code can open-source AI and allow a CNN to be scaled to other vehicle types. Additionally, the car was able to learn a course simply from image recognition. CNN can be user-friendly for broader use; however, unclear image recognition can skew data-processing and hurt the efficiency of RL. The details of description will be presented in the poster.

Poster 5008

Real-time Face Recognition Using Deep Learning

Cheng Zhu

Department of Computer Science
Felician University, Lodi, NJ

Faculty Advisor

Sofya Poger
Felician University, Lodi, NJ

During the past few years, deep Convolutional Neural Networks (CNNs) have achieved promising results in face recognition. The state-of-the-art performance of face recognition using deep learning techniques combined with large datasets have surpassed human level performance on the Labeled in the Wild (LFW) benchmark [1][2]. Today, there is a huge demand for efficient real-time face recognition systems on mobile devices due to increasing mobile use in human life. For a mobile device, a face recognition system needs a tradeoff between accuracy and speed under constraints of its computational power. In this project, we deployed a real-time face recognition system on Raspberry Pi 4 using deep learning techniques from OpenFace [3]. The system implemented a conventional face detection method, histogram of oriented gradients (HOG) with linear support vector machines (SVM), to detect human faces. The detected faces were fed to a deep neural network pre-trained by OpenFace to compute 128-d face recognition embeddings as feature vectors for face classification. In this work, we explored the applicability of recent deep learning techniques to mobile devices with limited computational resources.

References

[1] Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 1701-1708.
[2] S. Balaban, “Deep Learning and Face Recognition: The State of the Art,” Biometric and Surveillance Technology for Human and Activity Identification XII (2015): n. pag. Crossref. Web.
[3] B. Amos, B. Ludwiczuk, M. Satyanarayanan, "Openface: A general-purpose face recognition library with mobile applications," CMU-CS-16-118, CMU School of Computer Science, Tech. Rep., 2016.

Poster 5009

Real-World Algorithms: Shortest Paths

Mark Whaley
Department of Computer Science
Felician University, Lodi, NJ

Faculty Advisor
Sofya Poger
Felician University, Lodi, NJ

The shortest path problem is one of the most common problems in algorithms. There are many algorithms that can calculate the shortest paths through a graph, but they all have their own drawbacks. To solve the problem of navigating from a starting point to any other point, we will consider the most general problem of finding all paths from a starting to any other point in our graph. In this work, we analyze the Dijkstra algorithm and the Bellman-Ford(-Moor) algorithms. We also describe the solution for the shortest path problem for planar graphs using the seven delta-wye-delta replacements.

The Dijkstra algorithm is greedy, and does not work with zero or negative weights. Bellman-Ford works with negative weights, but if all weights are positive, the Bellman-Ford will make many unnecessary calculations, increasing run time. Delta-Wye-Delta, however, seems to be the best of the algorithms. Rather than traversing graphs, Delta-Wye-Delta reduces them, making finding the shortest path not only simpler, but also more efficient.

We also present real-world applications for shortest-path problems: Internet routing and arbitrage. In internet routing, we apply shortest-path estimates for nodes in graphs that are directly connected to our starting node, so these will be the shortest paths containing only one edge. We repeat this procedure, and relax again along all edges of the graph. After |V| - 1 iterations of the relaxation process for all edges we find the shortest path from our starting node to every other node in the graph. In arbitrage we detect negative weight cycles for detection of arbitrage opportunities.

Poster 5010

Predicting Capture-to-Control Delay in Automated UAV Systems

Rahul Thapa

Department of Computing Science
Villanova University, Villanova, PA

Faculty Advisors

Marco Levorato
University of California – Irvine, Irvine, CA

Benjamin Mitchell
Villanova University, Villanova, PA

Edge computing has vastly mitigated the limitations of onboard hardware on Unmanned Aerial Vehicles (UAVs). However, the lag between sensing and control in such a system, referred to as capture-to-control delay, still poses a problem [1]. The performance of such system can be drastically improved if the delay can be modeled and predicted. We studied the feature space of such a UAV system which included Telemetry as well as Network data and selected features that have a high correlation with the delay. We hypothesized that the current state of the system might predict the delay that happens in the future. Therefore, to model that dependency, we used a Recurrent Neural Network with LSTM units to model the delay as a function of the selected features. For our experiments, we collected data using a 3DR solo drone**.** We then used various tools such as scikit-learn, minepy, MIQ, and MID based approaches to explore the feature space. We finally selected the 10 best features from the set of around 100 available features**.** For modeling, we used a network with 3 LSTM layers followed by a fully connected layer using a softmax activation function as output. Across our experiments, the model was able to capture the anomalies/delays (in one experiment, as high as 70%) despite having minimal training data and highly skewed classes. We found that our model failed to predict sudden short duration delays but did a good job predicting prolonged delay. We even found that is some cases, our model was able to predict the delay a few time steps before it happened.

References

[1] D. Callegaro, S. Baidya and M. Levorato, "A Measurement Study on Edge Computing for Autonomous UAVs," *Proceedings of the ACM SIGCOMM 2019 Workshop on Mobile AirGround Edge Computing, Systems, Networks, and Applications*, pp. 29-35, 2019.

Poster 5011

Computational Topology in Music Analysis

Khalid Glenn

Department of Computer Science
Felician University, Lodi, NJ

Faculty Advisors

Sofya Poger
Felician University, Lodi, NJ

We proudly present my research in Computational Topology in respect to Music Analysis. In simple terms topology is the modern version of geometry, the study of many shapes and curves in multi dimensions. What separates topology from geometry is that any shape that is presented, even if the shape may inquire deformities, in topological terms the shape is still considered the same. Our original topic of study was graph theory with its pairwise relationship with objects, then we stumbled across a subdivision of graph theory "Topological Graph Theory" which studies immersions of graphs. Eventually this led me to the field of Topology. Topology has a subdivision that indulges in the applications topological techniques called Topological Data Analyzes or TDA. Topological Data Analyzes contains many different interesting fields of study such as data mining, computer vision, and Computational Topology in Music Analysis. "Computational Topology in Music Analysis" is the application into how musical data can be looked at by analyzing the sound it's wave of music and then converting them into what is known as Tonnetz which is a conceptual lattice diagram representing tonal space. Then by using Topological techniques the data is then represented in the form of Simplicial complex which is a set composed of points, line segments, triangles, and their n-dimensional counterparts. This Simplicial complex gives analyzers a better and more efficient way of looking at the musical data and musical Analysis. Our poster will discuss in detail different ways to analysis musical data in comparison with utilizing topological techniques, it will also show future applications of using topology to understand data.

Poster 5012

Major Miner

Xin Liu
Computer Science
Brooklyn College, Brooklyn, NY

Faculty Advisor
Michael Mandel
Brooklyn College, Brooklyn, NY

Professor Michael Mandel and his team are building computer systems to emulate humans' ability to identify sounds even in the presence of noise. Building such machines requires training data in the form of human descriptions of sounds. Major Miner is a web-based sound labeling game that collects "tags" (descriptions) from humans to better train these machines. The rules of the game were designed to encourage players to provide reasonable and creative descriptions of sound clips, or single 10-second segments of a sound recording.

This game was initially built in 2008 using Ruby on Rails 2.0 and collected descriptions of music. The goal of the current project was to re-implement this game in a modern web framework and apply it to recordings of environmental sounds from Alaska. Previous students working on this project had almost completed the user interface components using the JavaScript framework React.js, but they had not implemented the gameplay functionality.

Thus, we implemented this gameplay functionality, which required writing the entire back-end of the system in Firebase in addition to making major modifications to several user interface components. One major piece of this was building the entire scoring system, allowing players to score immediately by matching another player's descriptions or have the potential to score more points later when a subsequent listener matched their description. We made the game more competitive by creating weekly and daily leader boards. The selection of clips was implemented to keep both new and experienced players engaged. We improved the performance of the website by optimizing the algorithms and implementation. And finally, we deployed it to a production server and configuration.

Poster 5013

Choose the Best Airbnb Host in NYC: A Geo-Visualization Project

Mohammed Alzuwayyid

Computer Information Systems
SUNY Buffalo State, Buffalo, NY

Faculty Advisors

Sarbani Banerjee
SUNY Buffalo State, Buffalo, NY

This research project will focus on analyzing the datasets from Airbnb of New York City (NYC) to answer the following questions. Which hosts are the busiest and why? Is there any noticeable difference in Airbnb reservations in various areas of NYC, like Queens, Brooklyn, Manhattan etc.? What are the factors that are attributable to that difference? Is there an increase in the reservation between different seasons? What can we learn about different Airbnb hosts located in the NYC areas? This project will be developed using Python programming language to do the cleaning, normalizing, visualizing and analyzing of the dataset. Jupyter Notebook will be used as IDE as well as other libraries like Pandas and NumPy, Matplotib and GeoPandas. Furthermore, a website with HTML5 coding will be developed to visually display the data and the map. The poster presentation will be about different areas of NYC where the busiest hosts are located, and it will also show the reasons why guests choose those hosts, alongside predictions and comparisons about the prices and locations. A website will be developed for this project to show the map of NYC with the places where most reservations are made. It will also allow the users to explore different details such as review per month, prices, room types and reasons for reservations about the hosts who are the busiest. This will help the user make educated decisions about their choice of Airbnb hosts in NYC.

Poster 5014

Skip the Line and Get Your Coffee on Time

Bandar Almutairi

Computer Information Systems

SUNY Buffalo State, Buffalo, NY

Faculty Advisors

Sarbani Banerjee

SUNY Buffalo State, Buffalo, NY

The purpose of this research project is to design an electronic Order Application (App) system for Spot coffee at the Student Union, which is located on the Buffalo State Campus. The coffee shop often has long queues of students during the morning hours and break times between classes. The large crowd of students gathers before and between classes; they have very little time to pay through the traditional cash register system, which currently is the only payment method available. This project will automate currently the ordering system and offer cashless payments to reduce the staff needed to take orders and payments. The project is aimed to develop the App for both Android and iOS smartphones so that students can order their coffee and make online payments. This project will also create an App for coffee shop through which the order can be processed by the coffee shop staff. With this App, the baristas will be able to visualize the order requests and efficiently service the orders to the customers as they come in. An "Order Ready" message will pop up on the screens of customers' smartphones. The project will be expanded to create a website ordering system as well. The payment system will be based on PayPal and credit card payments.

Poster 5015

Using an Internet Extension to Save you From the Internet

Yousef Abdulrahman

Computer Information Systems

SUNY Buffalo State, Buffalo, NY

Faculty Advisors

Sarbani Banerjee

SUNY Buffalo State, Buffalo, NY

In recent times the Internet has become a major part of everyone's lives, and it can absorb one into its endless abyss of information and entertainment. Some people have become obsessed with the Internet, sacrificing their health, time with friends and family, and even becoming addicted. The purpose of this research project is to develop a Google Chrome Extension that monitors the time spent browsing on websites like Google, YouTube etc. It will be using JavaScript, HTML and CSS. First, the user interface will be created using CSS to have a user-friendly view of the Extension. JavaScript will be the main programming language used to implement the functionality of the Chrome Extension. It will show warnings every 30 minutes and to remind the user to take a break or change posture or hydrate. It will also send daily, weekly, and monthly statistics of the improvement on time spent on the Internet versus time spent on useful, constructive activities. For example, how many hours spent watching Netflix or YouTube compared to how many hours on educational websites like Buffalo State Blackboard. The web browser extension will be an icon at the top of the user's browser and will be hidden while the user is on the Internet; a pop-up window on the computer will show up on the screen to tell the user to take a break with an inspirational quote to get the user moving. This Extension will be published in the Chrome Web Store for anyone using the Google Chrome web browser to use.

Poster 5016

ReverseU

Steve Bumbera, Mathew Piscitelli, Ilya Shavrov, Adnan Sumra, Tyler Thomas
ECECS
University of New Haven, West Haven, CT

Faculty Advisor
Ibrahim Baggili, Liberty Page
University of New Haven, West Haven, CT

We present ReverseU in the ECECS department at the University of New Haven. The goal of this project is to create a collaborative framework environment for reverse engineering. There is a shortage of effective reverse engineering resources tailored towards beginners and as a result, a shortage of qualified candidates in the workforce that can respond to current national security concerns. We believe that our work will not only encourage newcomers into the world of reverse engineering but also serve as a platform that industry professionals can use. The main tool utilized for this project is Ghidra: a Software Reverse Engineering (SRE) suite of tools developed by NSA's Research Directorate, as the main disassembler and decompiler. To start, we examined the Ghidra source code and developed a custom API in order to call Ghidra functionality. Next, we redesigned Ghidra's graphical user interface in order to work on a web interface. Finally, we planned and developed the infrastructure needed for the environment. Our poster will discuss our reasoning for using Ghidra, the design decisions made and their respective justification, and the challenges we encountered during this project. The full description and implementation features of the framework will be presented in detail.

.

Poster 5017

Land Maverick

Aidan Conahan, Ryan Morton, Taliyah Rivera, Jenna Uba, and Austin Webber

Tagliatela College of Engineering
University of New Haven, West Haven, CT

Faculty Advisors

Liberty Page
University of New Haven, West Haven, CT
Emily Yale, Larry Studwell
Land Maverick LLC, Brandford, CT

The goal of the Land Maverick project is for the team to replace the software of an existing robot so that it can be controlled by a remote as well as run autonomously. The purpose of this robot is to monitor the health of soil on large pieces of property, initially targeting golf courses, by using a navigation system and measuring sensors. The team is adding data collection software which will verify when the robot is stopped, deploy the measuring sensors, save the measurement data, and retract the measuring device. The data collection software will then communicate with the navigation software to resume traversal. The team created navigation software will use highly sophisticated algorithms (using a customer created algorithm) in order to traverse the course and reach the target sample spots. The team uses data from various detection sensors such as GPS, Infrared, and Time of Flight, will work alongside the traversing algorithms to ensure the robot avoids hazards such as sand traps and water sources. The data collection software will be written in python and the navigation software will be written in C++.

Poster 5018

Confluence of industrial experience and academic knowledge to enhance students' self perception of future employability

Zachary Maxwell, James Heenan

Department of Computing Sciences
SUNY Brockport, Brockport, NY

Faculty Advisor

Jerry Ajay, Sandeep Mitra
SUNY Brockport, Brockport, NY

A key factor that associates an employer with a candidate is previous work/internship experience in their related field. Job/Internship opportunities can be intimidating for students as some may convince themselves that they don't have the experience to apply.

Towards this end, we conducted a study wherein we had a local industrial partner for whom we developed a front-end GUI using Vue and a back-end database utilizing MySQL. The application involved development of a tool that enables migrating spreadsheets into an SQL database.

We noted that in the course of this real-world project, when the student development team interacted with industrial personnel, and were regularly evaluated by them, the experience was significantly different from a typical academic one. We report three key findings that we acquired through this experience.

First, in academia, students submit assignments, visit office hours, are part of a rigid structure with well-defined assignments and have a professor who has all the answers. During an industrial project, a student does not necessarily have these advantages. Consequently, we learned that academic projects are like learning to swim with floats whereas industrial projects are similar to learning to swim while thrown into the deep ocean.

Secondly, we learned that accumulating experience with technologies desired by employers (such as Vue/Kotlin) on our own was important for showing motivation outside the classroom. Such contemporary technologies are not necessarily taught in school.

Thirdly, through industry experiences, we learned that students can develop strong relationships with individuals in the field, which is great for career networking.

With these insights, we conclude that industrial experiences teach new practices unexposed to students in academia. Industrial projects should not be intimidating to students but, rather, despite the uncertainties in requirements and technical challenges involved, should be viewed as an opportunity to enhance one's self perception of future employability.

Poster 5019

Comparison of Brain Activity with Agglomerative Clustering

Alexander Avery
Computer Science
Stonehill College, Easton, MA

Faculty Advisor
J. Timothy Balint
Stonehill College, Easton, MA

Neuroplasticity plays a vital role in the development and mastery of new tasks. In mammals such as mice, we can detect the influence of environmental change on the function of the hippocampus. Environmental enrichment increases neurotrophic factors and gives rise to incredible improvement on hippocampus-dependent learning. In many other studies related to neuroplasticity, the measurement of new brain activity usually stops at the level of activations in different regions of the brain. By harnessing agglomerative clustering on EEG datasets, we can build a behavioral pattern language for brain activity. The clustering model ran on data representing how each brain region is activating, as opposed to merely focusing on what regions of the brain are active. Varying the number of clusters in the agglomerative clustering algorithm brought the model closer to an accurate representation of the different kinds of thought patterns the participants used. The participants from the dataset were completing tasks for which they are untrained. The clustering model utilized the EEG dataset from the University of California, San Diego, in which participants attempted to recognize animals in an image quickly. The expectation was that clustering the brain activity of the participants would group participants with similar accuracies. The final clustering model shows that for the animal recognition task, different patterns of brain activation give rise to little difference in a participant's accuracy for the given task.

Poster 5020

Automotive Brand Perception Using Twitter Sentiment Analysis

Sydney Calvert
Mathematical Sciences
Eastern Connecticut State University, Willimantic, CT

Faculty Advisor
Garrett Dancik
Eastern Connecticut State University, Willimantic, CT

Buying a vehicle is a great expense for many people. With so many options to choose from, it can be useful for potential customers to know what others think or their experiences with that company. Twitter has become a social media platform in which millions of users can share (tweet) their opinions in 280 characters or less. Although many companies have Twitter accounts, it can be rather difficult to analyze the thousands of tweets that mention their company name. Sentiment analysis is a natural language processing technique that infers emotions, opinions, and attitudes from written language in an efficient way.

The primary objective is to compare the sentiment of tweets about various car manufacturers, however traditional sentiment analysis methods have known limitations when used on tweets, which can be sarcastic and can contain pictures and emojis. I therefore first set out to compare the performance of sentiment analysis methods. Specifically, I compare Naïve Bayes and the *R* package *sentimentR*, which contains various dictionaries, to assess how effective the methods are for determining the sentiment of Twitter data. Twitter datasets were obtained from Figure Eight, a platform that provides data with crowdsourced sentiment scores. I then used the most accurate method to analyze tweets regarding the best-selling automotive brands in the United States, according to CarFax. My poster will discuss the performance of the various methods and give insight as to whether tweets are predictive of car sales.

Poster 5021

Centennial Robot

Justin Frannis, John Lang, Amber Marrero, Me'Lia Ramos, Noah Jorgensen, Brendon Malouf, Charisma Banks-Obanor
Tagliatela College of Engineering
University of New Haven, West Haven, CT

Faculty Advisor
Liberty Page, Vahid Behzaden, Marya Neary
University of New Haven, West Haven, CT

In celebration of the University of New Haven turning 100 years old, we are creating a robot to showcase what the University of New Haven has grown to accomplish since its inception. The goal is to advance the students Artificial intelligence abilities with real world experience through helping with the client had envisioned for the Robot. The robot will be able to recognize important figures to the university, converse with those people, either limited or fluent, navigate a given environment autonomously, and look presentable to the guests. The facial recognition utilizes Tensorflow to train the artificial intelligence model while conversation will use IBM Watson. Autonomous navigation is implemented by using the Turtlebot, open source API and the Turtlebot 2i which is used as the base of the robot. This project has given us a chance to showcase what we have learned in our computer science classes as we had to use our coding and problem solving backgrounds to implement and make a functional robot. Using machine learning and artificial intelligence, our robot will navigate within an area, recognize faces, make small talk, and answer questions regarding the University's history with attendees/alumni at the Centennial Ball. This robot will be used in future University of New Haven publicity events to exhibit the capabilities of the College of Engineering.

Poster 5022

Roll Your Own Crypto: Use Cases for Novel Encryption Algorithms

Joshua Stover

Computer Information Systems
SUNY Buffalo State, Buffalo, NY

Faculty Advisor

Sarbani Banerjee
SUNY Buffalo State, Buffalo, NY

With data breaches on the rise and the looming threat of vulnerabilities in established encryption algorithms inching closer, data security is a growing need. Security must be considered at every step in a piece of software's lifespan, and in every component therein. This research project will employ a novel message-digest algorithm that will generate 512-bit, 256-bit, or 128-bit keys from a user selected passphrase. The message-digest algorithm will use a call to the system time to salt the hashes it generates, ensuring the uniqueness of keys even if passphrases are similar. In addition to the message-digest algorithm, it will implement an unique symmetric-key encryption algorithm using the aforementioned key to encrypt and decrypt files. The software will be written in C++ and compiled for use in Windows 10 and Ubuntu Linux. A simple interface for each platform will allow file selection and choice of encryption level. The project will be created using Visual Studio Code in conjunction with Qt Creator. The presentation will show machines running the software, encrypting and decrypting data from local and removable disks. The poster will examine the mathematics and the code used by the algorithms, as well as an exploration of use cases. The project aims to show that creating novel encryption algorithms can increase the data security of an organization or an individual without being any more cumbersome than implementing an existing algorithm.

Poster 5023

Practice, Assessment, and Review: The Pedagogical Model Behind an Intelligent Tutoring System

Kerry Buckman
Computer Science
Ithaca College, Ithaca, NY

Faculty Advisor
Toby Dragon
Ithaca College, Ithaca, NY

I present my experience participating in the Summer Scholars Research program with the Computer Science Department at Ithaca College. The focus of my research was to design a pedagogical model: an automated teaching model influenced by the theory and practice of learning. This model is used in an Intelligent Tutoring System (ITS) for Cornell University's College of Veterinary Medicine. The ITS, named PAR: The Practice, Assessment and Review System, allows students to check their knowledge and practice their skills related to equine ultrasounds. Pedagogical models use knowledge about students to make informed decisions on what questions to present. The PAR pedagogical model I developed considers students' knowledge levels based on responses to previous questions to decide which questions to display. In PAR, different types of questions have different difficulty levels, and the student must answer a certain amount of questions within a window to move onto harder ones. There are also follow up questions that will appear only if the preceding question is answered correctly. The repetition of questions is minimized, avoiding memorization of answers and creating a more comprehensive experience. I was particularly intrigued by this portion of the ITS because as an Education Studies minor, I have a strong interest in teaching, learning, and related technologies. This spring, the larger project team will pilot the PAR system in courses and conduct a study regarding the effectiveness of the system, including the pedagogical model. Participants will evaluate the speed at which PAR is presenting new kinds of questions and the appropriateness of their difficulty levels. My poster will include a visualization of how the pedagogical model interacts with the rest of the system, an in-depth description of the pedagogical model, details on our study, and a visual of PAR's interface.

Poster 5024

Using Oracle APEX to Solve Meetings Issues

Yasir Alhadeethi

Computer Information Systems
SUNY Buffalo State, Buffalo, NY

Faculty Advisor

Sarbani Banerjee
SUNY Buffalo State, Buffalo, NY

In this research project, Oracle Application Express (APEX) is used. APEX is a low-code development platform that enables one to build scalable, secure enterprise apps, with world-class features, which can be deployed anywhere. Oracle Application Express enables you to design, develop, and deploy responsive database-driven applications, either on-premises or in the cloud. Corporate meetings often can be overwhelming when it lacks documentation. Searching for files for each meeting can take up unnecessary time due to lack of organization. This project will provide a better environment for efficient collaborations and better facilitation among project team members. The APEX Application for this research is used to store data from meetings, including agenda, meeting organizer, assigned presenters, and other attendees. This application will contain specific details like name, description, start and end times, location, presenters, attendees, attachments, action items, decisions, and notes. The user can click a tab and the data about the meetings will be e-mailed to participants using an Oracle client. This project will use a database to store details and information that will be entered and created by the users of the application. This action will be deployed by using database objects like Tables, Views, Indexes, SQL Queries, Shell Scripts, as well as SQL Scripts. This application will prevent users from losing documentations and missing important details of each meeting, and thereby enhance communication.

Poster 5025

Use of GraphViz in VSCode to Dynamically Visualize Adjacency List/Matrix Based Graphs

Jack Cutri
Computer Science
Ithaca College, Ithaca, NY

Faculty Advisor
Ali Erkan
Ithaca College, Ithaca, NY

Graphs are one of the most fundamental data structures taught in computer science education. They serve as abstractions of countless types of networks (ranging from technological structures to metabolic pathways) and create context for impactful algorithms such as Dijkstra's and Ford Fulkerson's. Consequently, their thorough coverage is typically part of one or more SLOs in data-structures/algorithms courses. The implementation of graphs also affords us an opportunity to emphasize the important principle of programming-to-the interface: much like lists can be implemented with arrays vs linked lists and maps can be implemented with balanced trees vs hash tables, graphs can be implemented with adjacency lists vs adjacency matrices.

Computer Science education literature contains a multitude of reports that focus on how graphs and graph algorithms can be visualized for effective instruction. While our work is situated in this realm, our approach is based on the hypothesis that the impact of visualizations is particularly deep if they are built by the students themselves.

In this poster, we outline our use of the open source tree/graph visualization software GraphViz in the context of VSCode's GraphViz extension. In particular, we outline our implementation of a set of Java classes that allow graphs to be rendered in real-time within VSCode. These visualizations (which can themselves be built by students) depict the difference between the adjacency lists and adjacency matrices implementations of graphs. Since they can be updated by the running of a program (especially in the context of a debugger), our visualizations also provide insight about the operation of graph-based algorithms.

Poster 5026

Deconstructing Discord

John Day

Computer Science & IT Department
Chestnut Hill College, Philadelphia, PA

Faculty Advisor

Pam King
Chestnut Hill College, Philadelphia, PA

The goal of this research was to determine what digital artifacts can be found on a MacOS computer that had been using the Discord desktop app. During a summer internship at the software company, Sumuri an experiment was preformed to test what artifacts were present. The experiment consisted of the creation of a new test Discord account and using this account to create a server that would be populated with various types of messages. After the server was salted with data the next step was to find what information was stored locally by the Discord app. Once locally stored data was found the next step was analysis. The process of analysis involved using programs that permit the identification of what parts of the operating system is being affected when Discord is running. Once these files were identified the final step was to examine what information was found in them and determine if they were forensically relevant.

What was found in the analysis process was a cache file that had saved much of the activity that took place in the test server locally. Data of interest in this cache file included various forms of plain text files containing messages and image files that were sent to and from the test account. Overall the Discord cache proved to be a great repository for information regarding a user's activity on the platform.

Poster 5027

Blockchain Technology Adoption on College Campuses

Ali Al Mahmoud

Computer Science Department

Chestnut Hill College, Philadelphia, PA

Faculty Advisors

Lisa M. Olivieri

Chestnut Hill College, Philadelphia, PA

Blockchain has developed beyond being simply a financial instrument, since its inception. Its primary focus is cryptocurrencies.Th e associated value proposition, as it continues to grow and provide new functionality, has had a great impact on the companies and organizations utilizing it.

This study explores the reasons why cashless transaction using blockchain has become part of everyday use by assessing the relationship between blockchain and its underlying value drivers. Key areas analyzed and discussed are decentralized applications and transactions as well as ethical considerations and issues surrounding blockchain technology adoption. Additionally, this study addresses effective management issues related to blockchain adoption.

With blockchain standing out as one of the most secure and faster payment systems, its adoption within college campuses is explored as part of this study. The blockchain project named “Griffin Coin” was developed with the aim to make it easier for college students to make online purchases inside college campuses. The “Griffin Coin” project consisted of an Android mobile application wallet, back-end Blockchain client-server that monitored the movement of transactions and a marketplace website where students made online purchases.

The results of the project are as follows: The client generated wallets and viewed the transactions made in the network. The marketplace website communicated with the blockchain back-end, made deductions from the wallet, and sent transactions to the blockchain. The Android mobile application successfully communicated with the blockchain back-end and marketplace website.

Poster 5028

The Effects of k-Opt Algorithms on the Efficiency of the Snakes and Ladders Heuristic for Finding Hamiltonian Cycles in Graphs

Kat Cannon-MacMartin
Mathematics and Computer Science
Marlboro College, Marlboro, VT

Faculty Advisor
Jim Mahoney, Matt Ollis
Marlboro College, Marlboro, VT

I present my research on the effects of using standard k-opt algorithms in conjunction with the Snakes and Ladders heuristic (SLH) proposed by Baniasadi, Ejov, Filar, Haythorpe, and Rossomakhine in their paper *Deterministic "Snakes and Ladders" Heuristic for the Hamiltonian Cycle Problem*. As noted by Baniasadi et al., replacing the first stage of SLH with an efficient k-opt algorithm has the potential to improve the performance of the heuristic. My research consisted of experimenting with various previously-developed k-opt algorithms to investigate how they compare to SLH. The Hamiltonian Cycle Problem (HCP) is one of the most significant problems in mathematics, and is known to be NP-complete. SLH tackles this problem by repeatedly searching for Hamiltonian cycles with an upper-bound on computation steps, forcing the heuristic to finish in polynomial time. To determine if the performance of SLH could be improved, I implemented a selection of standard k-opt algorithms to replace the first stage of the heuristic. I used step-counting to measure performance in order to eliminate various uncontrolled variables. As anticipated, results were mixed, with some algorithms performing generally better than the first stage of SLH and others performing generally worse. My subsequent analysis of the data includes performance comparisons across a variety of tested graphs, as well as projections of how the tested algorithms would affect the overall performance of SLH when used in place of the first stage. My poster will provide background information on HCP, as well as a breakdown of SLH and information on the alternative k-opt algorithms that were tested. The poster will also include a summary of my testing process, and the resulting data will be presented along with detailed analysis.

References:

Baniasadi, Pouya et al., *Deterministic "Snakes and Ladders" Heuristic for the Hamiltonian Cycle Problem.* Mathematical Programming Computation 6(1), 55-75, 2014.

Poster 5029

Drone Forensics: The Right Drone for the Job

Elena Ortiz

Computer Science & IT Department
Chestnut Hill College, Philadelphia, PA

Faculty Advisor

Pamela King
Chestnut Hill College, Philadelphia, PA

With new technologies, the use of drones for government (CIA, FBI, and military) and digita orensics use is part of our everyday lives. The goal of my research was to compare and contras he similarities and differences of data extracted and analyzed from an expensive DJI Spark Drone and the inexpensive Protocol Director Drone. The experiment consisted of flying both lrones using the Samsung Galaxy S8 Active phone as a remote controller and comparing the data hat was extracted using Cellebrite software (UFED 4PC and UFED Physical Analyzer).

The results indicated that the DJI Spark Drone was much more advanced than the Protoco Director Drone. It would be a better choice to use in more advanced operations. The data generated, extracted, and analyzed in this experiment can be used as a resource in future cases because the more expensive drone was proven to be a lot easier to use. Finding the right drone fo he job will make it easier for the military, law enforcement agencies, and enthusiasts to execute heir intended operations.

Poster 5030

Incognito Mode Artifacts Found in RAM

Amy Nguyen
Computer Science & IT Department
Chestnut Hill College, Philadelphia, PA

Faculty Advisor
Pamela King
Chestnut Hill College, Philadelphia, PA

Being one of the most popular web browsers, many computer users take advantage of Google Chrome's simplicity and speed. With its own private browsing feature, users can also utilize Incognito mode. Incognito mode claims not to save browsing history, cookies, cache, and form data. The goal of this research project was to discover private browsing artifacts in RAM that would have significance during an investigation. Specifically, Google Chrome and Incognito mode were used. GoSecure's RAM acquisition (FDPro) and analysis (ResponderPro) tools were used to recover residual Incognito mode data. Browsing artifacts were created with a wiped Windows 10 system in Google Chrome's standard browser and Incognito mode to compare the discovery of data between both browsers. The standard browser and Incognito mode's memory dumps resulted in containing the same amount of data. As a result, browsing artifacts in both the standard browser and private browser were easily found using memory forensics.

Digital forensic investigators identify malicious behavior using residual data found in RAM by performing acquisition and analysis procedures with memory forensic tools. The results of the project indicates that if a digital forensic investigator captured a system's random access memory (RAM) that uses a private browsing mode, they could find potential artifacts for evidentiary value.

Poster 5031

Self-Driving Car

Ares Canetti, Andrew Walsh, Nick Kovalchik

Computer Science
Merrimack College, North Andover, MA

Faculty Advisor

Christopher Stuetzle
Merrimack College, North Andover, MA

Our team has created a self-driving car that is capable of using machine learning to gather data on certain images, so it can recognize and react to various situations. The goal of the project was to create a car that was capable of using machine learning to guide its actions, and to not have to rely on direct human control. Since commercial self-driving cars are already being tested for consumer use, it won't be long until they are a reality. This makes it very interesting to work on a very basic version of that idea. The car is able to use an attached camera to collect pictures of the oncoming road, and compare those against data learned from training the car on pictures of stop signs and obstacles. Using this method, the car can check pictures of a road to recognize stoplights and obstacles in that picture. From here, it can react according to what it detected. It is also able to recognize the lane lines on a road and use those to turn and stay on track. Our poster will show the development process and internal workings of our application, as well as go into more depth on the individual features of the car.

Poster 5032

Rove Photo

Kevin Benkart, Ares Canetti,Nicholas Paolino and Emma Theberge

Computer Science and Visual Performing Arts
Merrimack College, North Andover, MA

Faculty Advisor

Christopher Stuetzle, Dan Vlahos
Merrimack College, North Andover, MA

We developed an iOS application for sharing geotagged photography and photogenic locations among both amateur and professional photographers as a joint independent study between the Computer Science and Visual and Performing Arts departments at Merrimack College. ROVE, as we would come to call it, was a solution to a problem one of our members noticed in their field of photography, which was that there were no resources to share, search and find photoshoot locations. And so we created an iOS application that allowed users to place pins(markers of locations on a virtual map) in places they've visited, tag them to help others find it, and place photos of the location; to help other users find the perfect places for there photography needs. We developed the app in ReactNatve using Expo and used Firebase to store generated information. Locations are stored in Firebase as a coordinate pair, along with a Super pin that describes its location type (coastal, rural, city etc.), and sub pin types that describe notable characteristics of the site (graffiti, abandoned, architecture etc.). During the building of this app, we ran into technological challenges, such as adding assets into the expo build and creating a storage system for organizing data in Firebase such that it could be easily written to and read by our application.

Poster 5033

Application of Data Mining to Drug-Drug Interaction

Steven Runyon, Giuseppe Guerini, Jake Collier, Leah Starnes
Computer Science
Stockton University, Galloway, NJ

Faculty Advisors
Duo (Helen) Wei
Stockton University, Galloway, NJ

This project is rooted in the domain of healthcare, but more specifically concerns the dispensation of both over-the-counter and prescription drugs. Often, certain drugs cannot be recommended or prescribed to a patient because of allergens or other health complications that the drug is not designed to treat; in these cases, an alternative drug is required. The research surrounding this project not only hopes to discover better methods to identifying those alternative drugs but hopes to provide general insight to medical professionals on what additional drugs they can recommend to their patients. The methods to identify candidate drugs and Drug-Drug Interaction (DDI) that are exercized by this particular project are co-occurrence and association rule mining.

We first collected data pertaining to heart disease from ClinicalTrials.gov, a public database composed of over 300,000 clinical trials. That data was then preprocessed using natural language processing (NLP) methods in order to extract drug names from each clinical trial and map them to standard drug classifications. For our analysis, we developed two separate algorithms using concepts of co-occurrence and association rule mining. Our co-occurrence mining algorithm produced a matrix of co-occurences between any pair of drugs and a list of the drugs that co-occur with each other most frequently. Our association rule mining algoritm, based upon the Apriori algorithm, produced a list of drug to drug association rules. Finally, we analyzed the results and presented some of the findings.

Our goal was to identify the statistical relationships between drugs used throughout clinical trials. Several significant relationships were found. For example, ticagrelor and clopidogrel are significantly related as they co-occurr frequently and serve similar purposes.

This project was funded by the Stockton Research and Professional Development (R&PD) grant.

Poster 5034

Analysis of Transport Layer Congestion Control in 5G Networks

Quan Chau

Department of Mathematics and Computer Science
Dickinson College, Carlisle, PA

Faculty Advisors

Farhan Siddiqui
Dickinson College, Carlisle, PA

The Millimeter-Wave (mmWave) wireless technology has the potential to provide very high data rates and is expected to become a key enabler of 5G communication. However, mmWave signals suffer with high penetration loss and poor isotropic propagation which can result in intermittent packet losses. Transmission Control Protocol (TCP), which is the Internet's de facto transport protocol employs Congestion Control (CC) algorithms that consider packet loss as an implicit notification of network congestion and react by reducing the data transmission rate. This project focuses on analyzing how TCP performance is impacted by its congestion control techniques when data is transmitted over mmWave radio links. I selected various TCP versions (NewReno, YeAH, Hybla, WestWood, and Vegas) to determine how the CC algorithms impact the achievable data rate. The CC algorithms supported by these TCP versions differ in their reaction when packet losses and delays are experienced. I simulated the TCP versions in different scenarios based on the 3rd Generation Partnership Project (3GPP) requirements and standards, which covers a wide range of realistic settings such as dense urban areas, indoor hotspots, and rural areas. I use network simulator 3 (ns-3), a discrete-event network simulator for Internet systems. With the data collected, I focus on analyzing the performance of each congestion control algorithm when the user device moves through Line-Of-Sight (LOS) Non-Line-Of-Sight (NLOS) conditions. Our poster will discuss our motivation for evaluating TCP performance over mmWave links and I will present detailed results about each TCP version based on metrics such as data transmission rate, throughput, Round Trip Time (RTT), and Signal-to-Interference-plus-Noise Ratio (SINR).

Poster 5035

Exploring Deep Learning Techniques for Sound Source Separation in Multimodal Audio-Visual Context

Abi Pradhan, Kunal Sheth
Department of Computer Science
Connecticut College, New London, CT

Faculty Advisor
Ozgur Izmirli
Connecticut College, New London, CT

Sound source separation is a difficult problem to solve. Previous methods have adopted mathematical and statistical approaches to solve this problem. Given current advancements with deep learning architectures and computational speed, we aim to explore if machine learning can give us better results. Some of the popular architectures used currently are UNets and ResNets. Previous deep learning methods have only leveraged a single modality - sound. Our approach revolves around using multiple modalities - sound and vision - to help us better understand and process the input. We chose the University of Rochester Multimodal Music Dataset (URMP) which is a collection of 44 classical music compositions ranging from duets to quintets where each piece has its mix track along with individual instrument tracks. Each piece also has a video of the musicians playing those instruments.

We use two different architectures to process information from the individual modalities. We use YOLO for computer vision and WaveUNets for sound separation. The input (video file) is first fed to YOLO which outputs the list of instruments being played in the video. Then, the sound from the video is passed into the WaveUNet to get the individual tracks as outputs which are conditioned on the vision module's output.

We trained 13 different WaveUNets models, one for each target instrument in the URMP dataset. We also trained a YOLO network to detect these 13 instruments from video. Our main motivation to use a multimodal approach was to compare how using information from the visual domain changes the accuracy of the audio source separation. We want to evaluate whether this does a better job than the blind source separation approaches that have been tried out so far. Our results are comparable to the state-of-the-art results that have performed separation on URMP dataset.

Poster 5036

Improving the Quinnipiac Weather Website

Ryan Hanlon, J. Eamon Duffy
Software Engineering
Quinnipiac University, Hamden, CT

Faculty Advisor
Stefan Christov, Kimberly DiGiovanni
Quinnipiac University, Hamden, CT

Project Description: The project focuses on collecting real-time weather data from sensors within and near Quinnipiac University and visualizing this data via public website. The website also provides the city of New Haven a dashboard with the data from the city's sensors. We plan to contribute the weather data we collect to a crowdsourced website, such as Weather Underground.

Background/Contribution: Our team has taken over a previous team's web application and has re-written it in a modern web framework to improve the application's maintainability. We also added the functionality to obtain weather data collected outside Quinnipiac (i.e., from New Haven weather sensors) and display that data. We further enhanced the project through adding the capability to crowdsource the collected weather data.

Methods: We used the Scrum[1] methodology to develop the web application. The web application's front-end and back-end use the React.js[2] and Express.js[3] frameworks, respectively. The data is stored using MongoDB[4] and is hosted on a local Linux server. The web application obtains data from the weather sensors via HTTP requests. The web application is evaluated by obtaining frequent feedback from our clients, advisors, and peers. We plan to administer a user survey to further evaluate the user interface.

Results: We built a website capable of collecting data from various weather sensors and displaying it. The website provides local weather data that is sometimes more accurate than weather data from public weather services based on relatively distant weather sensors. By storing the collected data, the site also serves as an educational resource for studying local weather patterns. Prior to our website, the city of New Haven was unable to easily access the data from some of their weather sensors without a third-party licensing subscription.

References:
1. Scrum Methodology: https://www.scrum.org
2. React.js: https://reactjs.org
3. Express.js : https://expressjs.com
4. MongoDB: https://mongodb.com/

Poster 5037

LibreFoodPantry: Diving into DevOps by Developing Open Source Project Infrastructure

Christopher Radkowski

Computer Science
Worcester State University, Worcester, MA

Faculty Advisor

Karl R. Wurst
Worcester State University, Worcester, MA

My research was focused on improving and developing the infrastructure and workflow documentation for the LibreFoodPantry community. LibreFoodPantry (LFP) is a faculty-led, humanitarian, free, open source software community which is developing software to support on-campus food pantries. I choose to do this research as I wanted to explore the type of work that someone with a career in DevOps would do, which is a career path I am interested in pursuing after graduation.

My research goal was to find the best tool to develop the LibreFoodPantry projects with. I also tested a proposed workflow out with these different tools to see how well it worked on each platform.

The LibreFoodPantry project is distinct in that faculty members from multiple institutions created an open source project that other faculty members outside of these institutions can use for their computer science courses to introduce students to developing for an open source project.

My methods consisted of researching the features of the different tools the faculty members selected to investigate and comparing them to see which ones are better.

Ultimately, a decision was made to use GitLab Gold as the tool to host and develop the LFP projects with. I helped migrate the projects to GitLab and I helped create documentation that detailed for future developers how to use the proposed workflow for working on the LFP projects.

Poster 5038

Python3 Simulation of Sugar Metabolism in Cancer Cells

Kenneth Paul Callahan

Natural Sciences
Marlboro College, Marlboro, VT

Faculty Advisor

Todd Smith, Matt Ollis, Jim Mahoney
Marlboro College, Marlboro, VT

Herein I present a mathematical model and accompanying Python3 simulation of altered glucose metabolism in cancer cells. During the progression of cancer, transformed cells undergo a metabolic switch to glucose fermentation for large scale energy production, which is extremely inefficient. To evaluate metabolite concentrations during glycolysis and fermentation, I created a model of 13 first order partial differential equations based on Michaelis Menten enzyme kinetics, a widely used model for single enzyme kinetics. All differential equations can be numerically solved with a user friendly Python 3.6.10 program that uses Euler's Method to estimate the concentration of metabolites over time. When run, the program provides background on the model, allows the user to manipulate the initial conditions and parameter values, and numerically simulates the concentration changes. The program holds data for "default" initial conditions and parameter values, and if the user elects to use their own, the program will display the default values for each parameter for reference. After the parameter values are set, a time series function will estimate the trajectory of each metabolite concentration for a user defined number of steps. These data can be written to a CSV file using a homemade CSV module, per user request. This CSV module can also import data from CSV files as a dictionary of lists, allowing the user to generate time series data and plot them at another time. The main product of this program are integral curves which estimate metabolite concentrations over time, given a set of initial conditions. Plots are presented using a Matplotlib and NumPy based plotting function, and allow the user to observe the system over time. Overall, this program shows the power of Python3 for numerically solving large systems of differential equations, and managing the data generated during simulations.

Poster 5039

TweeTeX: a Programming Language for Hypertext Literature

Nick Creel

Computer Science
Marlboro College, Marlboro, VT

Faculty Advisor

Jim Mahoney
Marlboro College, Marlboro, VT

I present my research in programming language and compiler design completed as part of my Plan of Concentration in Computer Science and Writing at Marlboro College. The goal of my research is to design a programming language with the functionality of other domain-specific programming languages designed for the creation of hypertext literature, such as the Twee programming language used by the popular hypertext creation software Twine, while also improving the syntax of the language so that the syntax is readable, clear, and intuitive. I model the syntax of my programming language after the TeX syntax introduced by Donald Knuth in 1978, and that new language, which I call TweeTeX, is cross-compiled to Twee3 source code. To design and implement the compiler, I create a grammar specification and design a compiler for my new language using a lexer, recursive descent parser, and code generator without using a lexer generator or parser generator. My poster will discuss these different components of the TweeTeX compiler in detail, in addition to the challenges I faced during design and implementation. I will also demonstrate the functionality of the language in detail with proof-of-concept examples.

Poster 5040

Off-Market Real Estate Investing: Exploring Data Analytics

Evinn Wheeler

Computer Information Systems
SUNY Buffalo State, Buffalo, NY

Faculty Advisor

Sarbani Banerjee
SUNY Buffalo State, Buffalo, NY

There are many real estate investors looking to invest in off-market properties. Many investors have trouble trying to maintain their current properties as well as finding new good deals all at once. This research project will help them to search for these deals and give them all the information they need to secure a great deal. American homeowners go through foreclosures and lose their homes every day. Instead of them losing everything, the current project will be able to connect them with investors to sell their home and at the same time will give the investors an opportunity to make a great return on investment. This project will be built with Python programming language. The Python modules used in this project Pandas, NumPy, Matplotlib, and Altair. The data will be coming from several database sources including the Buffalo city, Erie County Surrogate, Zillow, and Prop Stream Data. These databases contain data related to, property and tax records, mortgage information, foreclosure list, probate list, high equity list, projected home prices, the actual repair value of the home, and more. The project then cross-references all those lists and values to find out the likelihood of a homeowner facing foreclosure, possible sales price range, and the profit percentage of each home. This web-based project will be accessed by investors through the Google cloud platform. The poster presentation will include key aspects of the project as well as a live presentation of the software.

Poster 5041

Handover Mechanisms in 5G Millimeter Wave Networks

Sajal Satyal
Computer Science
Dickinson College, Carlisle, PA

Faculty Advisor
Farhan Siddiqui
Dickinson College, Carlisle, PA

My project aims to evaluate the performance of different handover algorithms for Millimeter Wave (mmWave) 5G networks based on factors such as handover latency, handover frequency, and network throughput. The assessment is performed by using a network simulator called ns-3 to mimic real-world environments where such networks will be deployed. My work involved the setting up and running simulations with relevant parameters and processing the data to do produce an insightful analysis of available handover techniques. In 5G networks, radio waves in the frequency range 30-300 GHz are used. These waves, enable higher data rates but deteriorate over short distances and cannot pass through obstacles. This behavior of millimeter waves necessitates frequent switching (called a handover) of cellular base stations by the mobile device to maintain a continuous network connection. Handovers should be minimized, since no data can be transmitted over the network during a handover. My work involves comparing three different algorithms: threshold, fixedTTT and dynamicTTT. When using the threshold algorithm, the network switches base stations whenever a neighboring base station offers a higher signal strength. Networks adopting the FixedTTT algorithm wait a predetermined amount of time to make sure that the neighboring base station still offers a higher signal strength, and only then perform the handover. DynamicTTT algorithm behaves similarly to FixedTTT except the time interval to re-check the handover criterion is calculated based on the difference in signal strengths between the current and neighboring base stations. My current research indicates that in an urban environment the DynamicTTT algorithm yields the highest throughput and lowest handoff frequency.

Poster 5042

Applying Mindfulness Techniques to a Smartphone App to Manage Screen Time

Zoe Beals

Computer Science
Skidmore College, Saratoga Springs, NY

Faculty Advisor

Aarathi Prasad
Skidmore College, Saratoga Springs, NY

We present our experience designing a smartphone app to help a user manage their screen time. The goal of this tool is to allow the user to become aware of their smartphone usage and distinguish between mindful and mindless usage. The motivation for tackling this problem is rooted in finding a better way to make a user self-aware of their phone usage without causing the stress and anxiety current screen time gadgets can cause. Practicing mindfulness is proven to have a positive effect on mental health (Shapiro et al). Combining the practice of mindfulness with screen time management can result in a beneficial effect on a user's relationship with their smartphone. This research ties together the practice of mindfulness and reducing the stress surrounding screen time management. Our smartphone app is intended to run in the background of a smartphone (similar to Apple's ScreenTime tool). We will present our design of the prototype smartphone app as well as our plans for conducting the user study to test the prototype app at the poster session.

References:

Shapiro, Shauna L., et al. "Mechanisms of Mindfulness." *Journal of Clinical Psychology*, vol. 62, no. 3, Mar. 2006, pp. 373–386., doi:10.1002/jclp.20237.

Poster 5043

Have Music Lyrics Changed Over Time? Analyzing Hit Songs Using Natural Language Processing

Jafet Aparicio

Computer Science

Eastern Connecticut State University, Willimantic, CT

Faculty Advisor

Garrett Dancik

Eastern Connecticut State University, Willimantic, CT

In the United States, the *Billboard* Hot 100 is the standard record chart used in the music industry. Songs are rated based on sales, radio play, and online streaming. As an aspiring artist, releasing a song and landing a spot in the *Billboard* Hot 100 is a career changer that can set one up for success. By carefully dissecting the songs that end up becoming hits, we can begin to understand the characteristics of hit songs in order to create a new song that has a high probability of being a hit. My main objective is to use natural language processing (NLP) in order to compare hit songs to non-hit songs based on their lyrics. NLP involves extracting and interpreting human (natural) language. Sentiment analysis involves computationally determining whether a piece of writing is positive, negative, or neutral. Lyrics were obtained through web scraping by using the *LyricsGenius* python package which scrapes song lyrics from the Genius website. Songs that reached the top 50 of the *Billboard* Hot 100 year-end chart since 1990 were considered hits, while songs in the bottom 50 were considered non-hits. Hit songs were compared to non-hit songs based on sentiment analysis, lyrical repetitiveness, word usage, and named entity frequencies. We also evaluate whether these features change over time. These findings will form the basis of future research that aims to predict hit songs from their lyrics as well as their musical elements.

Poster 5044

Improving Acronym Searches on PubMed

Kevin Williams

Computer Science

Eastern Connecticut State University, Willimantic, CT

Faculty Advisor

Garrett Dancik

Eastern Connecticut State University, Willimantic, CT

PubMed, an online repository containing >30 million citations of biomedical literature, is widely used by researchers and students. Currently, some PubMed searches are automatically mapped to Medical Subject Headings (MeSH terms) in order to improve coverage. MeSH terms are a controlled vocabulary derived by biomedical experts that are used for categorizing articles. Unfortunately, however, many common acronyms are not mapped to MeSH terms, and these acronym searches will yield potentially incomplete results. For example, a search for 'tumor necrosis factor' using the search term 'TNF' would only execute a query for an exact match to 'TNF'; however, a search for 'tumor necrosis factor' would map to the corresponding MeSH term and the search will account for alternate spellings ('tumour') and wording ('factor-alpha'). The latter is likely the intended search, which returns >30,000 more articles. In this research we develop a web browser extension that will allow researchers to better perform PubMed searches using biologically meaningful acronyms. In this work, we use Python to analyze >30 million PubMed abstracts in order to identify acronyms and their associated phrases, and then link acronyms to MeSH terms where possible. Finally, we develop a web plug-in which leverages the acronym-MeSH mappings so that acronym searches are expanded to include associated MeSH terms. When using the web extension, researchers will be able to find articles that mention the acronym, as well as have the associated MeSH term. This web extension will allow for more meaningful searches than the current PubMed searching methods. The poster will describe the methods used to identify acronyms from PubMed abstracts, to identify acronym to MeSH relationships, and document the development of the web plugin; as well as results demonstrating the use of the enhanced searching method.

Poster 5045

Multiplicative Agreement Edge Detection

Conner Saeli

Computer Information Systems
SUNY Buffalo State, Buffalo, NY

Faculty Advisor

Gang Hu, Sarbani Banerjee
SUNY Buffalo State, Buffalo, NY

The goal for this research project is to develop a novel edge detection method that would perform more accurately than existing approaches. This project originated from studying new edge detection methods using convolutional neural networks. Edge detection provides a global view of an image with the most critical outline of the image. Robust edge detection methods would provide more concrete analysis for computer-based image understanding. The applications for this are useful in a wide range of fields from satellite imaging to medical screening. For example, the results from medical images could be examined in batches to detect any anomalies without human intervention. Modern deep neural network-based edge detection methods achieve roughly 80% accuracy, so the intended result is to improve on this baseline. This model is based on a PyTorch implementation of a current edge detection algorithm called Holistically-Nested Edge Detection (HED). HED's approach is to calculate the loss and update its weights using not only each intermediate feature map but also with a single added result of each feature map called a fusion layer. Instead of using conventional methods of adding intermediate feature maps together from the deep network, this approach adopts multiplication into the network. The objective for this is to extract more important information from network layers than existing approaches. By multiplying instead of adding, this project aims to find agreement between intermediate feature maps instead of simply adding them together. This implementation of HED uses pre-trained parameters from the VGG-16 convolutional neural network as a starting point. The datasets used for this project are BSDS500 and NYUDv2, which are public datasets from UC Berkeley and New York University respectively.

Poster 5046

Convex Hull Click Emoji

Afia Semin

Computer Science Department
Skidmore College, Saratoga Springs, NY

Faculty Advisor

Aarathi Prasad
Skidmore College, Saratoga Springs, NY

Graphical passwords (GP) are more *usable* and *memorable* than text passwords (TP) [2], however GP is more susceptible to *shoulder surfing* (an attacker can observe what a user chooses as their password when it is an image as opposed to text). Convex Hull Click point (CHC) [2] was proposed as a usable, memorable and shoulder-surfing-resistant GP scheme, however users took about 70 seconds to login. During the setup phase of the CHC scheme, users have to select n number of images out of k images as their password; during the login phase, 3 out of the n selected images will appear on the login screen, which will create a triangle and the user has to click within this triangle for a successful login.

I designed a modified version of CHC called Convex Hull Click Emoji (CHCE) that takes advantage of the widespread use of emojis to improve login time while retaining usability and shoulder-surfing-resistance properties of CHC.

I conducted a theoretical security analysis that compares the proposed CHCE with other authentication schemes such as unlock patterns, PINs, TP, Persuasive cued-click point (PCCP) [1], and Passpoint [2] using the following factors: the size of password space, and the likelihood of different attacks including brute force, dictionary building and shoulder surfing.

I plan to conduct a two-week long user study to determine whether CHCE does improve the login time of CHC, while retaining the usability and shoulder-surfing-resistance properties. I will present findings from the user study as well as the theoretical security analysis in the poster.

1. Chiasson, Sonia, et al. "Persuasive cued click-points: Design, implementation, and evaluation of a knowledge-based authentication mechanism." *IEEE Transactions on Dependable and Secure Computing* 9.2 (2011): 222-235.

2. Gao, Haichang, et al. "A survey on the use of graphical passwords in security." *JSW* 8.7 (2013): 1678-1698.

Poster 5047

Fashion Image Classification with Deep Learning

Xin Tong
Department of Computing Sciences
Villanova University, Villanova, PA

Faculty Advisor
Benjamin Mitchell
Villanova University, Villanova, PA

I present my experience exploring the applications of Deep Learning algorithms to image classification problems in the Computing Sciences Department at Villanova University. The inspiration of the research is the increasing popularity of image recognition technology in the fashion and ecommerce industries, as well as my personal desire to organize my belongings by categories. The goal of the system is to be able to classify each input image of a fashion item and assign it to a specific label. Before building this model, I investigated image classification models for cats and dogs, and then decided that for my model I would use a Convolutional Neural Network (CNN) implemented with the Machine Learning library TensorFlow. The image databases used are fashion-MNIST and DeepFashion. After training, the average accuracy of classification for the each of the 10 categories used in this classification model is 90%. Potential applications of this fashion image classification model include but are not limited to: incorporated in a mobile application to organize users' fashion belongings, partnered with a recommender system to automatically generate outfit combinitions, and helping e-commerce platforms to identify users' target merchandise. Future work could include designing a more effective variation of CNN to increase the accuracy of classification model, and making the model more robust to ambiguous fashion images. The full description and implementation of the fashion image classification model will be presented in detail.

Poster 5048

Debugger Features and Recursive Algorithms Using Map-based Algorithm Visualizations

Tyler Gorman, Alissa Ronca, Zack Goodsell
Computer Science
Siena College, Loudonville, NY

Faculty Advisor
James Teresco
Siena College, Loudonville, NY

Our research involved expanding METAL's (Map-based Educational Tools for Algorithm Learning) interactive capabilities and algorithm visualizations (AV). METAL is a toolset for teachers and individuals to be able to show algorithms visually, while also having the ability to interact with the various algorithms. Before our work on the project, METAL allowed for users to see 7 different linear algorithms while also being able to change things like the speed at which the program ran. Even though the capabilities of the project before us were great, there was and still is much room to grow and develop it further. One of such advancements was the addition of recursive algorithms to the list of choices. Since JavaScript must run in a single thread, you cannot update the visuals of the screen while still in a for loop. Because of this, the way METAL is programmed is in a pseudo assembly-JavaScript fashion where for loops are jump statements. This means we cannot use conventional for loops (or recursive calls) with our code. So, when a user sees a recursive algorithm, we are producing it linearly in code. This requires us to keep track of all variable changes, graphic changes and so on (in short, all the details a normal recursive stack would). On top of this change, we have also added a breakpoint system to all our algorithms. A user can click on a line of code that is displayed on screen to have the algorithm stop when it is encountered. They can also set conditional breakpoints for METAL to stop at for the user to inspect what exactly is happening on screen, which vastly expands the learning/teaching capabilities of METAL. These conditions are dependent on what we allow, but things such as if statements, node values and more are allowed.

Poster 5049

Preprocessing Pipeline for Analysing Tweets about Mental Health

Karan Khullar, Randy Champagne, Brendan Hayek
Computer Science
Stockton University, Galloway, NJ

Faculty Advisor
Duo (Helen) Wei
Stockton University, Galloway, NJ

Mental illness is an important topic of discussion, which greatly impacts the overall health of the people it affects. One way to understand the mental illness community is by analyzing a mass number of tweets on the popular social media platform, Twitter. However, analyzing twitter data can be a daunting task due to the sheer size and complexity.

In this project, we proposed a cloud-based preprocessing pipeline that can facilitate constant fetching, compressing, organizing, and cleaning of the twitter data. This system fetches tweets pertaining to mental illnesses hourly, and then stores them in an organized fashion for future analysis. The tweets fetched are organized by the mental illness they are associated with, the type of analysis done to them, and the time that they were retrieved. This way, we can always look back at past results and see how they differ from the present.

The overarching goal of this project is to make a sustainable pipeline that can capture the essence of the discussion surrounding mental illness and how it changes over time. Once the initial system is in place, it will run for as long as needed in the cloud, and analysis will be as simple as writing another script and adding it to the scheduler. Examples of analysis that can be performed on these tweets include, sentiment analysis, data clustering, topic mining, and many more. We believe that making a scalable system that can be easily utilized to highlight the thoughts surrounding mental illnesses can expand professional knowledge as well as assist those directly affected with these illnesses.

Poster 5050

Audio Feature Extraction of Human Speech

James R. Young
Computer Science
Worcester State University, Worcester, MA

Faculty Advisor
Karl R. Wurst
Worcester State University, Worcester, MA

This project used Python to analyze audio samples of human speech and explore the best methods for doing so. Machine learning was used to attempt to duplicate past research on this subject and explore new potential insights.

Signal processing is the backbone of many cutting edge technologies that would not be possible without a solid understanding of audio signals and methods of processing them. Understanding these concepts is a prerequisite to analyzing human speech. While human language is complicated in its own right, there are many patterns found in an audio signal of human speech that can be discovered through signal processing and reveal new insights. These can be found using multiple feature extraction techniques and combining multiple features. Quality of human speech is subjective but there are common attributes that can negatively affect how a person is perceived, such as a having a stutter or speaking in a monotone manner. Efficiently detecting these traits and the extent to which they exist in a human voice could have possible clinical applications.

Python was used to analyze spoken audio samples, making use of libraries such as NumPy, Pandas, Matplotlib, and Tensorflow. There is extensive research concerning analyzing speech that goes back decades, and these libraries were used in an attempt to duplicate those results. Techniques were compared and contrasted with those in existing research. When necessary, open source datasets of recorded human speech were used. With supervised machine learning, models were trained to correctly identify features. Large datasets are required to properly train a model, so feature identification was limited. Still, the methods used to obtain these results can be used when larger datasets with more pertinent information is available.

Poster 5052

The Role of Blocks Programming in Forming Computational Identity

Emma Brousseau
Integrated Digital & Data Sciences
Emmanuel College, Boston, MA

Faculty Advisor
Mark Sherman
Emmanuel College, Boston, MA

My research is on the concept of Computational Identity (CI), and the role of blocks programming in the formation of it. First, I define computational identity to have three components: a user's perception that computing is useful, their feeling of self-efficacy using computing, and their feeling of belonging using computing. I provide rationale for the definition, then move onto providing evidence for the role of blocks in its formation. The research investigates what computational identity is, why it is significant, and if blocks contribute to it. Computational identity is emerging as an important feature to study when considering how to motivate students to continue engaging with computation. There has been limited prior research on CI, and none on the role of blocks programming in forming CI. To explore this, I first found literature on concepts related to the components of my definition, and applied previous research to show how increased identifications with those components has led to increased retention in computing. I also found studies on the well-tested benefits of blocks programming whose findings may demonstrate the importance and effectiveness of blocks programming in the development of CI. In the process, I have constructed a well-justified definition of CI, providing support for the claim that increased CI may increase retention of students in computing. My findings overall support the hypothesis that blocks programming is an important and effective tool in developing a student's CI. I am applying the results of this research to a survey intending to measure CI of students taking an introductory computing course, to be used Fall 2020. The work has opened up the area to new questions, such as what the effect of a strong, personally meaningful goal is in the development of CI and the impact of CI on students pursuing further education in computing.

Poster 5053

Crowdsourced Measurements of Internet Traffic Manipulation

Allison Turner
Computer Science
Wellesley College, Wellesley, MA

Faculty Advisor
Christine Bassem
Wellesley College, Wellesley, MA

The internet is a fairly new type of resource, the first message on ARPAnet only sent half a century ago. In that short lifespan, the internet has revolutionized politics, economics, social structures, and information as we know it. Various bodies have interests in controlling data flow on the internet, such as governments, internet service providers, content distribution networks, and content providers. These entities enforce their interests via censorship, differentiated service, routing policies, and other strategies, all of which have potential to affect efficiency and user experience. With this research, we seek to determine the control tactics implemented on real networks and how these tactics affect the user.

The effects experienced across autonomous system areas and borders are heterogeneous and complex. No testbed could match the medley of paradigms present in everyday consumer networks, and so it is from the public internet from which we seek to gather data. We propose to do so by distributing experiment software from a task crowdsourcing platform and then analyzing the measurements generated from the participants' computers. Ethical consideration has been given to the valuation of participants' labor, their liability of searching for certain content, and their privacy.

We are currently in the process of building the experiment software. The experiment software will record some details about the participant's computer, such as the capacity of their network interface card, the AS' executor, the connection provider, and the average bandwidth speeds available to the participant. The software will then run a body of network tests against domains of varying popularity and contentiousness to survey metrics such as request round-trip time, network hosts traversed in transit, and transmission capacity. These data will be analyzed for patterns in relation to the nature of the queried domain's content, business relationships between connection providers and AS executors, and more.

Poster 5054

Earthquake Data Visualization

Corey Domingos

Computer Science

Eastern Connecticut State University, Willimantic, CT

Faculty Advisor

Sarah Tasneem

Eastern Connecticut State University, Willimantic, CT

The purpose of this creative project is to visualize earthquake data using Python. This is important because it helps the general public and scientists to better visualize the dataset that was used. By mapping the Earthquake data set with Python it makes the raw data easier for the viewer to understand. Since raw data can be overwhelming to grasp, this Python software transforms data into a more comprehensible version for whomever is viewing. To accomplish this task a data set of recent Earthquakes needed to be acquired, which initially started with the local area of Puerto Rico. Then to make the data set more extensive it was expanded to the major earthquakes globally. In order to apply the data set it was necessary for an application software to be developed to help make the data set less intimidating for the viewer. To create this software many Python packages were implemented. Some of the major packages that were used in the creation of this application were Gmplot, Pandas, and Matplotlib. Although there are many Earthquake mapping applications offered on the internet, it became apparent that none of them offered any sort of comparison between the strength of the Earthquakes. In the application developed in this project, a bar chart was used in order to help the viewer visualize the strength differences between some of the major Earthquakes globally. By implementing this bar chart the viewer is able to fully grasp the true meaning of the data in a clear and concise manner.

Poster 5055

Simulation of Motor Impairment with "Reverse Angle Mouse" in a Head-Controlled Pointer Fitts' law Task

Mariah Papy, Duncan Calder, Ngu Dang, Aidan McLaughlin, Breanna Desrochers
Computer Science
Clark University, Worcester, MA

Faculty Advisor
John Magee
Clark University, Worcester, MA

We present our experience participating in a research project in the Computer Science Department at Clark University during the previous summer. The problem that our work addressed is the difficulty of developing a UI that is accessible for individuals with motor impairment (MI). Particularly, these individuals are not always available to assist with research or software development testing. Thus, we believed that Simulation software could both decrease the burden on MI individuals and help software designers gain some perspective on the difficulties that MI individuals face when using the software. To help address this issue, we proposed a system that simulates the challenges demonstrated by a motor-impaired user interacting with a head-controlled mouse pointer system. First, we developed an algorithm called "Reversed Angle Mouse" (RAM) that enables us to simulate an MI user's difficulty with small range precision by decreasing the user's control when narrowing in on a target. Then, we built the simulation program by embedding the Camera Mouse software and RAM together. Finally, through evaluation using Fitts' Law Task, a standard in assessing pointing device, we compared the throughput rate in bits per second and the percent error rate of our system between users with and without motor impairments. The statistical analysis of our experimental results was conducted using ANOVA software, and the results were statistically significant. The full description and implementation features of our research will be presented in detail.

Poster 5056

Visual Analytics using Tableau

Eunbi Kim
School of Computer Science
Kean University, Union, NJ

Faculty Advisor
Ching-Yu Huang
Kean University, Union, NJ

As data becomes more accessible, visualization methods are needed to help make it easier to understand the information. Analyzing and visualizing data helps researchers recognize the patterns of the datasets, understand the datasets without reading the values, and join multiple datasets together. Tableau is one of the most popular interactive data visualization software. By using Tableau, it is simple and easy to clean up datasets, find correlation, and create graphs such as geo chart, map chart, table chart, line chart, pie chart, and treemap chart. This project aims to study the correlation between gross domestic product (GDP) and human immunodeficiency virus (HIV). Large datasets related to the GDP and HIV were gathered from open data sources – the World Bank open data, Our World in Data and World population Review. The datasets were cleaned up with Tableau and Excel in spreadsheet format which is easier to manipulate the attributes. The raw datasets include some short explanations, titles, headers, or null values. It is also required to normalize the values for the same attributes between different datasets. For example, one dataset uses “US” in the country name attribute, and another dataset could use “United States”. Another example is one dataset uses “South Korea” and another dataset uses “Republic of Korea”. If the data is inconsistent, the statistical results could be wrong. Therefore, it is very important to clean and normalize the data before further analysis. After cleaning up the data, Tableau was used to visualize the joined datasets in graphs and find correlations. This project found a strong correlation between GDP and HIV.