

Proceedings of the Symposium for Virginia Tech  
Undergraduate Research in Computer Science



April 30, 2015  
4:00PM–6:30PM  
Owens Banquet Hall



## Table of contents for Research projects (R1-R19) and Capstone projects (C1-C7)

R1: Orion-X: Providing Enhanced Understanding of JavaScript Dynamism .....	4
R2: WalnutiQ: OO Model of Human Brain with 1 Theorized Learning Algorithm .....	5
R3: Stack Up.....	6
R4: StarSPIRE – Extension and Statistical Analysis of a Multi-Model Semantic Interaction Tool for Text Analytics.....	7
R5: AstrophysVis: Designing Intelligible Visualizations for Multi-Dimensional Astronomical Data .....	8
R6: Developing Mobile Android Exergaming Applications that Promote Physical Activity Levels in Adolescents .....	9
R7: Supporting Crowdsourcing Sensemaking in Big Data with Dynamic Context Slices.....	10
R8: Smartwatches for Kids: Apps for Learning and Fun .....	11
R9: Mobile M-CHAT-R/F: Autism Screening Research .....	12
R10: Designing EEG Displays for Non-Clinical Settings .....	13
R11: Collusion Attack Implementation by Retrofitting Android Apps .....	14
R12: MindMapper: Analyzing Context for EEG Data.....	15
R13: Symbolic Regression of Symmetries for Automated Scientific Discovery .....	16
R14: MindMapper: Presenting Actionable Information to EEG Users .....	17
R15: GPGPU Utilization on Mobile Devices.....	18



R16: Towards Usable Security: Interactive Visualization of High Dimensional Data for Program Anomaly Detection .....	19
R17: Computing k-Shortest Paths 60x Faster.....	20
R18: Hypergraphs: Algorithms, Implementations, and Applications.....	21
R19: Project and Task Management as a Meta-tool of Personal Information .....	22
C1: Pintos Symmetric Multi-Processing.....	25
C2: Implementing Efficient Unix-like Processes .....	26
C3: Securing Pintos.....	27
C4: MapTack.....	28
C5: Implementing Networking Functionality for the Pintos Operating System.....	29
C6: Pass It On! .....	30
C7: PintosGrind – Dynamic Memory Analysis on Kernels .....	31



# R1: Orion-X: Providing Enhanced Understanding of JavaScript Dynamism

Anuj Pareek, [anuj1@vt.edu](mailto:anuj1@vt.edu)

Eunjee Kim, [silverg@vt.edu](mailto:silverg@vt.edu)

Advisors: Shiyi Wei, Dr. Barbara G. Ryder

JavaScript is the lingua franca of Web programming. It is a flexible dynamic programming language that enables interactive client-side websites. Due to its popularity, it is desired to have advanced IDEs to support the development of JavaScript applications. Unfortunately, JavaScript dynamism poses great challenges for program understanding and it also renders automated software analysis techniques ineffective.

Therefore, the state-of-the-art IDEs are insufficient to support JavaScript development.

Our goal is to design advanced JavaScript IDE features using novel software analysis techniques. We have implemented a new feature that enhances the understanding of eval functions by summarizing the JavaScript code generated at runtime. We also aim to support advanced smart code completion as well as JavaScript object hierarchy visualization. These new features are implemented as plugins to extend Eclipse Orion, an open source cloud-based IDE.



# R2: WalnutiQ: OO Model of Human Brain with 1 Theorized Learning Algorithm

Quinn Liu, quinnliu@vt.edu

Advisor: Dr. Anthony Cate

Currently, the exact common learning algorithm and connection configuration of the neocortex (70% of the human brain) is still a mystery. WalnutiQ is an open source human brain model simulation in Java trying to solve this mystery with a foundation in biological evidence and mathematics. In it's current state WalnutiQ's API allows developers to build a generic model of layers 4 & 5 of the neocortex with a simple retina that is able to move. All code and work is hosted at [github.com/WalnutiQ/WalnutiQ](https://github.com/WalnutiQ/WalnutiQ). The long term goal of this project is to store code that can simulate a full sized human brain in real-time.



## R3: Stack Up

Koorosh Massoudi, km419@vt.edu

Bryce Langlotz, brycel@vt.edu

Hee Gu Yea, paulyea@vt.edu

T'Keyaha Sawyers, tks2015@vt.edu

Advisor: Dr. Osman Balci

The purpose of our independent study project was to implement a social, exercise tracking website to encourage people to exercise through competition. A main reason many people don't exercise is a lack of motivation; by allowing users to compare their progress through multiple quantifiable metrics we can encourage them to exercise more often. We set out to implement this competitive network in an easy to use way in an attempt to relieve the pain points of current exercise tracking mediums. The process we used was the incremental development model, allowing us to break down each component of the site into separate parts. Each portion of the site had a data storage and access implementation stage, a user interface implementation stage, and a testing and approval stage. This allowed us to simplify and evaluate each step more easily and efficiently. The end result is a more in depth approach to health and fitness than predecessors such as MyFitnessPal, which have minimal social aspects. Our site is a more communal health and fitness tracking application.



# R4: StarSPIRE – Extension and Statistical Analysis of a Multi-Model Semantic Interaction Tool for Text Analytics

Nathan Wycoff, nathw95@vt.edu

Katherine Gorbach, kate18@vt.edu

Advisors: Dr. Chris North, Dr. Leanna House, Lauren Bradel

StarSPIRE is an existing multi-model semantic interaction tool that enables users to explore vast document collections by inferring what information may be relevant to the user as well as where it should be arranged on the display while shielding users from the complexity of the underlying models and algorithms. This system was evaluated after gathering usability feedback through pilot studies and carefully designing a comparative user study to observe any differences between implicit and explicit querying (e.g. queries constructed based on user interaction) and explicit querying. The data collected from this study were statistically evaluated and visualized to infer the cause of weakly significant differences in performance and significant differences in externalization actions performed between the study conditions. Traditional searching proved to be ineffective, while complex system-generated queries (implicit and query by example) tended to provide more relevant information. Leveraging the results of this empirical evaluation, StarSPIRE was expanded to integrate web-based search engines, namely Bing and IEEE Digital Library. User interactions are passively converted into nuanced search queries. The results of these searches are parsed via entity extraction, ranked by search engine relevance, then ranked by StarSPIRE's current relevance profile based on extracted user interests, and finally placed in context of the existing spatialized documents. Through enabling real-time web-based retrieval, StarSPIRE is capable of allowing users to perform tasks such as investigative analysis or literature reviews, with the system and user co-creating a sub-set of documents that are perceived to be relevant. The implicitly formed queries strive to “fill in the gaps” of the user's knowledge and illuminate connections that may not have otherwise been detected. Thus, users are able to leverage and manipulate powerful back-end models in order to provide answers to real world questions.



# R5: AstrophysVis: Designing Intelligent Visualizations for Multi-Dimensional Astronomical Data

Mitchell Wagner, mitchw94@vt.edu

Advisor: Rachel Elliott, Dr. Chris North, Dr. Duncan Farrah

The spectral energy densities (SEDs) recorded for stars with varying stellar parameters (including surface gravity ( $\log g$ ), surface temperature ( $T_{\text{eff}}$ ), metallicity ( $Z$ ), as well as reddening due to dust between an observing telescope and star ( $E(B-V)$ )) are severely degenerate at certain wavelengths. Therefore, rigorous instrumentation design for photometry-based space science missions (which study astronomical objects through the electromagnetic radiation they emit) calls for careful photometric filter selection capable of breaking these degeneracies. Current filter selection methods consist of qualitatively assessing where to place filters based upon well-known spectral features. However, when faced with more than a thousand possible filter combinations and hundreds of thousands of different combinations of various stellar parameters, tremendous potential exists to revolutionize the process via computation.

We aim to optimize filter selection by quantitatively comparing the performance of a large number of filter sets based upon the accuracy of each set's simulated spectral energy distribution fitting on a synthetic spectral library generated from Kurucz models. Our project has explored the computational analysis of filters in this regard, and the graphical representation of that performance. We have used Python to generate several visualizations showing how accurately stars with given parameters were identified by individual filter sets, and investigated the best approach to making that information graphically accessible to researchers looking to spot trends in the data. Our goal is to develop a method that sees widespread adoption as the optimal technique for designing imaging instruments in telescopes.





# R6: Developing Mobile Android Exergaming Applications that Promote Physical Activity Levels in Adolescents

Sonika Singh, sonika06@vt.edu

Advisor: Dr. Scott McCrickard, Dr. Paul Estabrooks

Physical inactivity puts adolescents at risk for a number of chronic conditions, including obesity, diabetes, and hypertension, as they become adults. To promote physical activity (PA) in adolescents, research efforts have focused on exergaming approaches which combine exercise and gaming.

Past research indicates that PA levels significantly decrease from 88 minutes/day for children (6-11 years of age) to 33 minutes/day for adolescents (12-15 years of age). Further, adolescents from low socioeconomic status (SES), who are known to be less physically active compared to their middle-high SES counterparts, are among the highest users of mobile phones (77% of 12-17 years olds own mobile phones). Recently, Allen & colleagues demonstrated that mobile exergames significantly increased physical activity in adolescents in a feasibility trial.

Three smartphone based exergames were tested for middle school children and showed initial feasibility and effectiveness. The purpose of this project was to develop attractive mobile exergames for children and their families with smartphones. The project began with an improvement of graphic design and scoring systems on three games: Fish Out of Water, Color Hunt, and Space Rayders. These games were published to the Google Store and have been advertised through a variety of methods, including online forums, departmental listserv, and paper flyers. The number of downloads, frequency of play (session), and duration of play were recorded to determine the participants engagement. The findings of this study will be used to investigate the effectiveness of mobile exergaming to promote PA in children.



# R7: Supporting Crowdsourcing Sensemaking in Big Data with Dynamic Context Slices

Joe Fletcher, joevt@vt.edu

Zhizheng Chen, zc219@vt.edu

Vijay Kuruvilla, vijayk@vt.edu

Advisors: Dr. Kurt Luther, Dr. Chris North

The goal of this project is to investigate how crowdsourcing can solve difficult image analysis problems which cannot be solved by computer programs or humans acting alone. We leverage the strengths of scalable human intelligence by developing a system using paid contributors from online labor markets like Amazon Mechanical Turk ([www.mturk.com](http://www.mturk.com)). These “crowd workers” work on a smaller subset of the problem (referred to as “context slices”) and collaborate with one another and build off each others’ findings to find a preferred solution.

In our research, we developed a crowd sensemaking algorithm with context slices to identify a set of unknown images. In this system, crowd workers could support the search efforts of an individual researcher in real time. First, we collected a diverse sample of images to be used for the first round of testing. Next, we created a series of web pages with image identification tasks for crowdworkers using Mechanical Turk’s API. The first page displayed one of these images to the worker and asked them to (1) describe what they see, (2) identify the specific names of people, locations, objects, etc. that they recognize, (3) rate their confidence on a 0-100 scale, and (4) justify their confidence rating. Once this information is submitted by the crowd worker, the results are stored on our MySQL database. A second page generates a sensemaking loop by allowing workers to review other workers’ responses to the same image. Finally, a dashboard interface aggregates the results per image in real time, allowing the highest-rated answers to rise to the top.

We have tested our system with more than 50 workers from Mechanical Turk. Our preliminary results are promising. We find that even with minimal structure, workers identify a wide range of clues in the identification phase. Workers also use tools like search engines to dig up additional information to support their hunches. During the review phase, workers tend to reach consensus on which responses are best fairly quickly. Future work includes fostering deeper collaboration between crowd workers, finer-grained dynamic slicing of the image content, and use of visualization to interpret the information provided by workers.



# R8: Smartwatches for Kids: Apps for Learning and Fun

Sonika Singh, [sonika06@vt.edu](mailto:sonika06@vt.edu)

Advisors: Dr. Scott McCrickard, Shuo Niu

This research investigates how smartwatches can be a fun and engaging part of the technical education experience for 4th and 5th graders. The emergence of these new and inexpensive devices allows us to get technology in the hands of a large number of children.

Our research team developed apps for smartwatches and mobile phones that highlight the capabilities and challenges of mobile platforms. We developed an event for a computer club at Harding Avenue Elementary School that provided children with the opportunity to explore the watch apps. The event highlighted challenges related to smartwatches, including memory management, sensor programming, network issues, and display size. Future opportunities for this can be to explore the role of smartwatch in mobile development and programming education.



## R9: Mobile M-CHAT-R/F: Autism Screening Research

Ching Chen, chingyg@vt.edu

Kyle He, kyleh8@vt.edu

Harrison Wade, jameshwade@vt.edu

Advisors: Dr. Scott McCrickard, Dr. Angela Scarpa, Dr. Luke Achenie

This Toddler Screening app seeks to make available the M-CHAT-R/F screening test to a mobile device, where users will have a mobile experience that is more accessible and straightforward than the traditional pen-and-paper format. The app tries to preserve as much of the M-CHAT-R/F as possible—particularly the core 20 questions, follow-up questions, ordering of questions, scoring system, and demographic information—but with improved accessibility for those seeking to understand the questions and reflect on their answers. A core improvement centers on an automated scoring system that uses Machine Learning algorithms to more precisely identify trends in the questionnaire and demographic data. Converting the M-CHAT-R/F into a app on a mobile device means both doctors and patients will be using it, resulting in privacy concerns that are addressed through two app modes: doctor and parent/guardian, to keep a distinct boundary between the two types of users mentioned. Other features include Text-to-Speech, quick navigation, and ease of use. Deployment is planned in clinics in southwest Virginia when funding is obtained.



# R10: Designing EEG Displays for Non-Clinical Settings

Gabriella Jacobsen, gabzulu@vt.edu

Advisors: Dr. Scott McCrickard, Robert Beaton

The goal of this project was to create a series of interfaces demonstrating Electroencephalogram (EEG) readings in a more intuitive manner for the average, non-clinical user interested in EEG data. This resulted in creating a persona and storyboard about Tim Johnson, a 15 year old male high school student interested in improving his attention levels through EEG tracking and Cognitive Behavioral Therapy. Elements of the design include exploring the interactions between the design and Tim, and the design and Tim's parent, teacher, and psychologist. For example the psychologist is able to access many levels of Tim's EEG readings and filter them by activity so that she may better understand when and where Tim is most attentive. Our research effort focused on creating an engaging and intuitive system from which valuable insights based upon EEG data could be gained without unnecessarily interrupting the user's lifestyle.

Our research has created a design that can be accessed through a smart phone, smart watch or computer that is synced with a portable, small scale EEG such as the MindSet Device by NeuroSky. In our work we are developing a simpler, non-clinical method of viewing and understanding EEG data using principles of usability and universal design. This means we are focusing on developing a design that will not only solve the attention issues of our persona, Tim, but will also be capable of assisting other stakeholders hoping to measure and validate attention level improvement. We see future advancements in biometric sensing infographics and data display so that an average, non-clinical user may have a better understanding of their bodily conditions, and an ability to heal or improve. We seek to expand the possibilities of how EEG data may be perceived with the assistance of user interface and graphic design. Finally, as the capabilities of small scale portable EEG's improve; we are looking to explore the range of groups that may benefit from EEG capabilities and the information EEG's can provide.



# R11: Collusion Attack Implementation by Retrofitting Android Apps

Lance Chao, kisseki@vt.edu

Ke Tian, ketian@vt.edu

Advisor: Dr. Danfeng Yao

The past decade saw a burst in popularity of smartphone devices. Android currently holds the largest number of install-based devices with over 1.9 billion devices in use worldwide. However, users of Android devices are at risk of information theft, especially if they download applications from third-party app stores. Here I discuss a new model of android malware that involves delegating the malicious task over multiple infected applications. This works by exploiting a messaging mechanism within Android called an Intent that allows multiple applications to communicate with each other. Research has shown that such malware is undetectable by traditional antivirus scanners. For proof of concept, I provide a simple example of how two infected applications can work together for a malicious intent, and demonstrate that this malware can infect benign applications through software repackaging.



# R12: MindMapper: Analyzing Context for EEG Data

Jayanth Prathipati, jayanth@vt.edu

Advisors: Dr. Scott McCrickard, Robert Beaton

With the use of EEG devices, we can measure personal engagement and tailor this data so that users can identify daily attention trends and have it displayed to better suit their own needs. Mind Mapper was a look into how we could design and develop an app that optimizes this data and helps users reflect and make educated decisions and help improve their attention levels.

For this application, we used various design practices, like for example, creating personas to try to figure out what our target user really wanted and needed. We found that we needed to create each view in our application to help users get a compounded set of data rather than having the same data set in different forms.

The first view that we created was a map view. This was meant to be used like a heat map and show you the most attentive places that you been in for the entire day. This worked incredibly well to show users locational trends but failed longitudinally. We ran into problems where users had multiple events on the same day in the same location. This all fed into one category rather than multiple categories and ended up not bringing enough value to the user.

In order to combat that, we created a daily, weekly, and monthly view. The whole point of these views was to bring another set of context to the user. Going back to multiple events, these views made it easy to view the data in that context. That way users could view that they had different events from 1-2 and 4-5 in the same location and could interpret that data differently. In addition, this view made it easy to compare trends that two events shared, like for example, if they had the same mean or variance.

Finally, we made an At-A-Glance view so that users could easily see their current engagement. This makes it easier for them to get immediate feedback and it helps them understand how well they are doing. This could be used to see sudden jumps in attention, bringing an additional set of information to the user that would otherwise not be there.

In all, we designed a set of views based on what we thought were user needs. Building out these needs helped us build a better application as we were able to understand why users would want to use an application like this one.



# R13: Symbolic Regression of Symmetries for Automated Scientific Discovery

Joseph Mehr, mehr@vt.edu

Colin Shea-Blymyer, c0lin@vt.edu

Advisor: Dr. Benjamin Jantzen

We present an algorithm for automated scientific discovery that uses “dynamical symmetries”. A dynamical symmetry is a change in or intervention on a physical system that commutes with the time evolution of the system. The laws describing how a collection of causally related variables relate to one another is associated with a uniquely structured collection of dynamical symmetries. It has been proposed that searching for dynamical symmetries might offer an efficient way to discover scientific laws, since symmetries are generally easier to identify than the underlying causal relations. The aim of our project was to design and test an algorithm that implements the suggested procedure. We call the resultant algorithm EUGENE in honor of Eugene Wigner, whose work inspired symmetry approach to automated scientific discovery.

The first challenge any such algorithm faces is that the space of possible dynamical symmetries is uncountably large and so not directly searchable. We made the problem tractable by narrowing the space to the set of functions described by a context-free grammar that preserves the order of operation. The algorithm we produced harnesses a function class developed specifically to store functions as parse trees accordant to the grammar we defined. We further restricted the space of functions by bounding the search to functions that can be generated from terminal symbols in a fixed, relatively small number of steps. Even this bounded space of functions is large for even a modest number of allowed applications of the production rules, and is likely rich in local minima – functions that are close to being a dynamical symmetry. We used a genetic algorithm to allow us to search the reduced space with a favorable balance of complexity and accuracy. Specifically, we used symbolic regression. As opposed to classical regression, symbolic regression allows us to infer both the values of numerical parameters in a target function, as well as the symbolic form of the functions in the space that we’re searching. In this case, the genetic algorithm interacts with a system to test candidate expressions and eventually returns a set of parameterized function forms. Each such form is an infinite family of symmetry transformations that characterizes the causal structure of the system under investigation. In the future we aim to connect the system we’ve developed to real life systems, optimize our genetic algorithm, and implement optimizations of the execution of our algorithm as well.

Acknowledgements: The authors would like to thank Prof. Dennis Kafura for his assistance overseeing this work.





# R14: MindMapper: Presenting Actionable Information to EEG Users

Jayanth Prathipati, jayanth@vt.edu

Ching Chen, chingyg@vt.edu

Kyle He, kyleh8@vt.edu

Advisors: Dr. Scott McCrickard, Robert Beaton

Attention trends are hard to identify on current EEG devices. With the use of well-designed apps, we can present actionable and usable information to users so that they can reflect and improve their attention. Mind Mapper was a look into creating an application with views to help facilitate those specific needs. We ideated and got feedback from sets of users and performed an expert analysis in order to validate this application and make sure that it is a well-built machine.

For this semester, we used multiple EEG applications (use the sensors and try out each application in general). Based on our experiences we were able to make views with data that we found to be generally important to use. For example, having a summary view with trends that most people identified with came out of using these multitude of apps and being frustrated with being unable to identify any trends from them. We were able to come up with four screens from using this application, we came up with an annotations screen that showed you current attention levels as well as a way to log those current annotations with text, so that you could take a look at it later. We came up with a summary view that showed you small trends, such as mean, median, max, and min attention over the period of a day. We built out a map view to compare locational trends with attention. Finally, we built a graph view to view annotations as well as look at a day's worth of data globally. After building out these views, we put them through an expert analysis. We conducted a walkthrough in CS 5714 and gave each expert a set of heuristics to evaluate this application with. We used the feedback from that source along with having a usability expert from Google take a look at the walkthrough as well. This showed us that people wanted to compound this information and have a look at it in multiple ways quickly.

Using the expert analysis, we created a new set of wireframes and revamped development to fit those needs. In particular, we changed the transitions between each screen to make it significantly faster to go from view to view. That way, users could take this data and analyze it much faster. In all, we made an application based on an incredible amount of user feedback and we learned a lot of the needs of a user and how to meet those needs.



# R15: GPGPU Utilization on Mobile Devices

Jeff Robertson, thejar@vt.edu

Advisor: Dr. Eli Tilevich

The goal of this project was to prove the practicality of using general purpose graphics processing units (GPGPUs) for hardware acceleration on mobile devices such as smartphones and tablets for the purpose of increasing energy efficiency and improving battery life. Because GPGPUs are optimized for graphics calculations, they can also be used to for other applications that use similar calculations. This sharing increases performance and uses less energy than the same operations on a general purpose CPU alone. While this optimization is common in conventional environments to increase performance in graphically intense processes, with decreased power consumption merely being a side effect, it has not been used in mobile devices where power consumption is critical.

Many mobile devices have a basic GPU as part of their mobile “System on a Chip” (SoC). Even with these lower end GPGPUs, increased energy efficiency should still be achievable as long as the GPGPU supports the graphics programming interface OpenGL and its derivative OpenCL, which many SoCs do.

To test this strategy, I attempted compression and special effect processing on audio files, which are very calculation heavy processes. I ran the tests in an emulated environment using a CPU and GPU which, although not directly comparable to those in mobile devices, were nearly as powerful and energy hungry relative to each other as the processing units might be in a SoC. I used the standard android emulator and android SDK from Google to create the mobile device environment and Aparapi from AMD to do the hardware acceleration.

My preliminary findings are that processing a given audio sample using only the CPU increases power consumption by 10-15 watts, compared to less than 2 watts increase using the GPGPU. The GPGPU also takes only one third to one half as long. In other words, the use of GPGPU's on mobile devices in calculation heavy applications could result in substantial power savings and increase battery life.



# R16: Towards Usable Security: Interactive Visualization of High Dimensional Data for Program Anomaly Detection

Hannah Roth, hroth94@vt.edu

Advisors: Xiaokui Shu, Dr. Danfeng Yao

Anomaly detection is an important approach in cybersecurity that aims to model normal execution behaviors of programs and networked systems, and using these models to detect anomalies and intrusions. Typically, the automatic detection needs to be followed by manual screening of security analysts in order to further eliminate false positives (false alerts). Despite the recent advances in designing anomaly detection techniques, there has not been systematic research on how to design usable tools for security analysts to effectively conduct the manual inspection. The specific goal of this project is to design an interactive visualization tool for a clustering-based program anomaly detection system.

Clustering data points that represent function behaviors allows security analysts to determine anomalies based on the proximity of a point to a cluster. The data that has been collected is multidimensional with each dimension representing an event, or sequence of function calls, and the corresponding value is the amount of times that event occurs. A visualization of this data must be created in order to assist security analysts with anomaly detection. The challenge presented is how to build a visualization tool that preserves similarities of multidimensional points while reducing the dimensionality. Further, another problem arises from the fact that each data point has the potential to occupy a different subset of dimensions. Our tool will be useful for security analysts to efficiently screen and investigate anomaly detection results.



# R17: Computing k-Shortest Paths 60x Faster

Nicholas Sharp, nsharp3@vt.edu

Advisor: Dr. T.M. Murali

This work presents a new augmentation to Yen's k-shortest path algorithm which dramatically improves performance on practical problems. Yen's k-shortest path algorithm is used to find the k-shortest simple paths from a source to a target in a directed graph, here it is investigated in the context of the human protein signaling network.

Yen's algorithm is based on the property that the  $i$ 'th shortest path must be a deviation from one of the previous shortest paths. This property is used to find k shortest paths by making many calls to Dijkstra's algorithm as a subroutines, each over a subset of the input graph. In this work, we demonstrate that the A\* search algorithm can be effectively used in place of Dijkstra's algorithm for all input problems, without an external heuristic function. This is done by using the distances from the source in the input graph as a monotone heuristic function for A\* search, which can be precomputed with a single call to Dijkstra's algorithm.

We investigate Yen's algorithm in the context of biological signaling pathways. A signaling pathway is a subgraph of a protein interaction network which connects receptors (sources) to transcription factors (targets). We consider the important problem of constructing this subgraph given the sources and targets, and demonstrate that computing k-shortest paths is highly effective, although many ( $k=20,000$ ) paths must be used.

When applied to the signaling pathway problem, the A\* augmentation to Yen's algorithm improves performance by a factor of 60. An analysis which would have required 8 months of CPU time using Yen's original algorithm was performed in just 4 days.



# R18: Hypergraphs: Algorithms, Implementations, and Applications

Brendan Avent, [bavent@vt.edu](mailto:bavent@vt.edu)

Advisor: Dr. T.M Murali

This overarching goal of this project was to explore how hypergraphs, a generalization of graphs, can be used in real-world applications. This required first understanding both hypergraph theory and the important algorithms that have been established in the literature, as well as how hypergraphs have been used in the past. For real-world applications, understanding hypergraph algorithms on their own is useless; software is required to execute such algorithms. Since there seemed to be no software libraries that provide support for both hypergraph data structures and algorithms, we created one. `halp`, the “Hypergraph Algorithms Package”, is our open-source hypergraphs library (now publicly-available on GitHub) that solves this problem. It contains hypergraph data structures, usable algorithms, and helpful utilities to create powerful hypergraph models and then to operate on them as easily as possible. Finally, we used the library to model biological networks, and even created a new hypergraph algorithm using Mixed Integer Linear Programming to give new insight into the nature of biological signaling pathways.



# R19: Project and Task Management as a Meta-tool of Personal Information

Rachel Kohler, rkohler1@vt.edu

Advisor: Dr. Manuel Pérez-Quñones

The goal of this project was to study task and project management and address information fragmentation by integrating independent, existing tools and leveraging cloud storage and client/server architectures. Independent tools are well suited to their particular collections and allow a user to complete a specific task. Our approach allows for the continued use of these independent tools. However, these individual tools often contribute to information fragmentation and are not conducive to project management. To get work done, we often need an integration of multiple sources of information stored under different tools. Our meta-tool allows integration of those sources, while supporting multiple approaches to project management.

Our tool provides a simple, lightweight interface which allows users access to multiple sources of information such as to dos, Dropbox, Gmail, Google Calendar, etc. for the purpose of project management. Several views allow users to focus on different activities in the planning, organizing, and completion of projects and tasks. An overview displays high-level information of all projects as well as a high-level view of the user's calendar. The project view integrates information from to do, email, and files for a particular project. This view allows users to focus on a single project integrating all information from different tools that comprise the project. The priority view shows a high level view of information relevant to priority projects or to-dos including to dos and emails the user has marked as priority and upcoming tasks. This view allows users to focus on important tasks that may not necessarily be related to the same project. These multiple views support a variety of project managing styles in an effort to accommodate individual preferences.









# C1: Pintos Symmetric Multi-Processing

Andrei Marcu, andreim@vt.edu

Nathan McCloskey, npgm@vt.edu

Nathan Waggoner, nwagg@vt.edu

Advisor: Dr. Godmar Back

Pintos is a simple instructional operating system designed to teach undergraduates core principles in operating system design. However Pintos lacks support for symmetric multiprocessing, a feature that has existed on the x86 architecture since the early 1990's. SMP leverages multi-core computers so that they can run multiple user-level applications at the same time. In the context of the Systems Capstone class, we chose to extend Pintos to support symmetric multiprocessing. This will allow Pintos to run programs concurrently. In order to implement this functionality, it is necessary to implement the Intel Multi Processor Specification to boot code on individual CPUs, remove all uniprocessor assumptions in the Pintos kernel, implement spinlocks in order to protect shared resources across multiple CPUs, and finally modify the scheduler to have a shared queue. To complete these goals, we ported startup functionality from MIT's xv6 operating system, and modeled our solution after other open source operating systems such as Linux and openBSD.



## C2: Implementing Efficient Unix-like Processes

Brianna Beitzel, bribytz@vt.edu

Marc Evangelista, marcato@vt.edu

Scott Pruett, spruett3@vt.edu

Advisor: Dr. Godmar Back

The goal for our capstone project was to extend the existing Pintos operating system framework to support useful, high-level user programs. As given, Pintos is a functional platform that supports kernel level threads, as well as other basic operating system features. To extend Pintos in a meaningful way, we borrowed concepts from Unix such as pipes, shared file descriptors, and process forking. Together, these allow composition of small, modular user-space programs to produce an effective working environment. In our implementation, we used modern strategies such as atomic reference counting and copy-on-write virtual memory mappings to increase efficiency and reduce memory overhead when forking new processes. On top of the low-level system, we have ported various user libraries to support high-level programs. This includes both C I/O and math libraries, as well as software floating point support. For functional user processes, we have implemented an expandable heap along with a user-level implementation of malloc. The resulting system is complete enough to run a Unix-like shell with pipes and I/O redirection, as well as an interpreter for the Lua programming language.



## C3: Securing PintOS

Christopher Schweinhart, [schwein@vt.edu](mailto:schwein@vt.edu)

Patrick To, [pkt2@vt.edu](mailto:pkt2@vt.edu)

Richard Hart, [rhart92@vt.edu](mailto:rhart92@vt.edu)

Audrey Decker, [adecker@vt.edu](mailto:adecker@vt.edu)

Advisor: Dr. Godmar Back

PintOS initially had no notion of security whatsoever. There was no concept of a "user", much less authorization, so anybody could read, write, and execute any file they wanted, leaving the operating system extremely vulnerable to adversaries. Adding security is a crosscutting concern, affecting all parts of an operating system, from process management to the filesystem implementation. It also requires support for authentication as well as privilege escalation.

Our group has addressed these issues in order to implement security in PintOS. We chose to model our design after UNIX, in order to learn more about it. First, we defined a user/principal abstraction, which allowed us to map each process to an effective user ID. Then, we implemented access checks for resources such as files and directories. Finally, for authentication, we used the MD5 message-digest algorithm to transform any given user's plaintext password into a 32 digit hexadecimal hash, which is then stored in a file called `/etc/passwd`.

However, we recognize that since UNIX is a 40-year-old technology, it is not a perfect security model.



## C4: MapTack

T'Keyaha Sawyers, tks2015@vt.edu

Advisor: Dr. Osman Balci

In CS 4704's our goal was to learn cloud software development and to create a cloud software that utilizes this technology. We ended up building an application called MapTack. MapTack aims to solve the problem of users keeping track of where the people they follow and their pictures were taken and manually tagging them with those locations. By creating MapTack it grants users the ability to upload their photos and tack them to a specific location while viewing where the users they follow have also taking pictures. The cloud software features we implemented were RSS Feeds, Web Chat, EJBs (Database Persistence), and File uploading (specific to images). For RSS Feeds, we used a Spring Servlet that allowed us to push out the appropriate info to those who had followed other users. Our implementation of Web Chat was done by using Web Sockets. By using the Java EE platform, we took advantage of its Server Connection Pools and JDBC Connection Resources. It allowed us to link our database and server, create classes based upon the tables and relationships in the database, and persist that data when it changed. By using Primefaces, a Java Server Faces library, we were able to use its file upload component and another library allowed us to pull out the meta data we needed to place the pictures on the map. At the moment, there are no future plans to continue on. However, we may want to continue and expand for a native Android and iOS application.

(Project members include graduates, Jason Riddle, Ashley Rose, Byron Becker, Sean Rice, Scott Sines)



# C5: Implementing Networking Functionality for the Pintos Operating System

Samuel Maddock, samm@vt.edu

Thomas Dinkel, tdinkel@vt.edu

Antoine Dahan, adahan@vt.edu

Advisor: Dr. Godmar Back

Any viable kernel today needs a way to communicate with other networks, including the Internet. The current iteration of the Pintos instructional operating system doesn't provide such functionality. For this project, we've implemented a networking interface for user-mode applications to interact with using the Pintos kernel. With Pintos running on top of the QEMU machine emulator, we were able to target a single network card, the Intel E1000. We implemented a network driver for the E1000 card according to the Intel 8254x PCI specification, allowing for the transmission and reception of Ethernet packets. We integrated the open source lwIP network stack into Pintos for interpreting packets and to allow us to provide BSD-style system calls to user programs. In order to demonstrate the networking additions to Pintos, we created a simple web server and a remote shell, each of which utilize the complete network stack.



## C6: Pass It On!

Minahm Kim, minahm92@vt.edu

Matthew Chittum, mc2315@vt.edu

Josh White, josh7411@vt.edu

Advisor: Dr. Kurt Luther

Pass It On! is an online marketplace that will initiate a social change with an economic shift from monetary value to that of time. This system hopes to promote social exchanges while addressing the problem of lack of reciprocity as seen in favors.

Through our research, we found that there is a growing apprehension to the current currency, and an increasing number of people view it as unfair. This opinion stems from the wealth gap consistently growing between economic classes and “bubbles” in the economy that result in inflation. These factors, among others, have led to the popularization of alternative currencies throughout the United States, ranging from local currencies backed by the community to things such as time banks.

Built off the principle that the time of all individuals is equal, individuals are able to post their skills and also request the services of others at the face value of the time it takes to execute a service. Throughout the country there are an assortment of time banks; however, generally they are local and small scale which restricts the scope of their effective range. Our system borrows concepts from time banks while addressing the issues of expandability and consistency through the development of our system in a web format.

From our research we found that there is no product that is exactly doing what we are trying to do, which gives us a unique standpoint for our product. When completing market research, we looked into other services that were similar to our system and found that the closest parallel was that of time banks as seen throughout the country. However, these time banks were restricted to very limited regions and were conducted in ways that were not conducive to expansion.

During evaluation, the Pass It On! team will test for usability, and user confidence. The goals of this system are to provide a secure environment to exchange skills and services to build a stronger community, as well as provide an alternative to traditional currency. A less concrete goal of the system is to foster community engagement as well as social exchanges in a society that is becoming more and more rooted in technology and big corporation. Currently the group is looking into doing a local release in Blacksburg to help gather more user feedback and determine if a large-scale release and expansion is viable for Pass It On!



## C7: PintosGrind – Dynamic Memory Analysis on Kernels

David Feroli, [daferoli@vt.edu](mailto:daferoli@vt.edu)

Alan Cai , [alan8@vt.edu](mailto:alan8@vt.edu)

Kelsey Farenholtz, [kelseyif@vt.edu](mailto:kelseyif@vt.edu)

Advisor: Dr. Godmar Back

Our group's project is a simple version of Valgrind, a dynamic analysis tool that can check for memory leaks in a program, specifically made for the Pintos operating system that runs through the Bochs emulator.

We call this Pintosgrind. Our reason for creating this dynamic analysis tool came from a problem we were having while programming for one of our projects for our Systems and Networking Capstone. We were failing one test for our program, which, according to our professor, is usually caused by memory leaks. However, there were not any programs we could use to detect memory leaks.

Pintosgrind keeps track of memory allocations and memory accesses. When the kernel shuts down, we warn the user of memory accesses to places where data has not been written to or unallocated areas. We also warn the user of places in the user code where there were memory allocations that were not freed. With the aid of this program, we will be able to tell if our problem on the project for our capstone was caused by a memory leak.

