**"Structured Modeling for Biological System Inference"**

<table>
<thead>
<tr>
<th>Faculty Advisor: TBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mentor(s): Charles Swannack, III (MITRE)</td>
</tr>
<tr>
<td>Contact</td>
</tr>
<tr>
<td>Research Area(s): Statistical Inference, Data Structures, Modeling</td>
</tr>
</tbody>
</table>

**Description of project and expected outcomes:**

In engineering and the biological sciences we often seek to develop models to describe the behavior of a process or system in order to either to predict properties of the system behavior or infer aspects of how the system works. When the system is simple (or alternatively when only a few properties of the system are of interest) we often can develop an analytical model for the system with a few free parameters. Given a reasonable amount of data, one can perform a best fit for this set of free parameters so that the resulting model meets a prescribed error tolerance. However, when the system or process is the consequence of multiple complex interactions such as biological systems and/or computer networks typically one must take a much different approach and first learn the problem structure and/or the appropriate model before inferring the system or process parameters. This can lead to an exponential growth in the complexity of the algorithms performing the system inference and often one must iterate between the learning stage of the algorithm and the inference stage.

The complexity of the inference algorithm leads to an interesting modeling and design trade space as computational tractability.
must be considered. One can develop a high fidelity class of models for which fewer iterations between the model learning and parameter inference phases can occur or develop a lower fidelity and structured class of models which lends itself to fast inference algorithms. The first approach may yield superior results if suitable model parameters can be found. However, if the higher fidelity model cannot converge the later approach may yield superior results. Central to this trade space (and computational tractability) is the underlying computational architecture for the system where the algorithm is run.

**Prerequisites:** Motivated students with a strong background in mathematics, data structures as well as programming (Python, C or C++)

**The student on this task will be expected to:**
1) Learn the basic theory of constraint-based structured learning
2) Learn how to use the open source constraint-based learning tools (i.e. the bnlearn R package)
3) Learn the basic inference structures that map on to contemporary hardware architectures
4) Identify (with assistance) a biological process to model with available data sets
5) Implement the standard class model for this process
6) Augment the standard class model for this process so that it maps on to the basic inference structures that map on to contemporary hardware architectures
7) Perform and analysis on the performance of the augmented model
Assessing the Lytro LightField Camera for Biometric Imaging
Faculty Advisor: TBD
Mentor(s): Nicholas Orlans and Carl Picconatto (MITRE)
Contact:
Research Area(s): Biometric Identification Sciences, Statistical Inference, Data Structures, Modeling

Description of project and expected outcomes:

Lytro Illum [1] is a relatively new prosumer camera technology with imaging capabilities that potentially can enhance biometric imaging applications. Specifically, refocusing, enhanced depth of field, and wavelength filtering could offer significant improvements and easier combined acquisition for face and periocular images. There has been communication with Lytro engineers and partner developer programs and they have expressed interest.

The student will be expected to focus on the following tasks:
1. Learn how to use Lytro Development Kit (and optionally a dedicated prototyping board) for supporting custom imaging applications.
2. Establish a small series of near infrared targets and perform

The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.
sensor response assessments and calibration against those targets.
3. In conjunction with Lytro, determine if cutoff filter can be removed or modified.
4. Collect an instrumented set of face and periorbital images, and estimate martin-schultze eye color for both light and dark eyed individuals.

**Prerequisites:** Students should have an understanding of imaging, evaluating biometric systems, and programing in OpenCV/C++ and Python. Background work in computational image processing is desirable.


"**Probabilistic Reasoning in UAV Decentralized Tasking**"
Faculty Advisor: TBD
Mentor(s): Marin Halper (MITRE)
Contact :
Research Area(s): Artificial Intelligence, Computer Science

**Description:**
Unmanned Aerial Vehicles (UAVs) have become more prevalent and capable assets. Now teams of multiple UAVs can be deployed with disparate sensors to collect more complete information. MITRE has previously integrated a commercial-off-the-shelf UAV control system, employing an Android mobile device as the onboard mission controller and the open source Pixhawk as the
autopilot. Android apps can be developed for these devices to communicate situational awareness data between the assets in order to perform decentralized tasking. The team of UAVs then will reach consensus on the collective assignment of assets to tasks that maximizes the rewards for task completion.

The goal of this project is to develop algorithms that use scoring, planning and consensus techniques to implement decentralized tasking and intelligent information collection. Specifically, we are exploring different algorithms for calculating the score (or reward) for executing a task and for efficiently building all possible plans for asset-task assignment. Once the plans are built, we need to establish consensus on the set of plans that maximize overall reward. As part of our intelligent information collection, we are exploring probabilistic reasoning algorithms, such as Markov Decision Processes, for evaluating next steps for information collections. Finally, we are investigating the use of probabilistic reasoning to determine what new tasks should be executed by the system when collecting information from multiple sources.

There are several opportunities to get involved in this project. The candidate can develop algorithms for the scoring, planning and consensus used in decentralized tasking. Another possibility is to develop algorithms for the probabilistic reasoning used for the intelligent information collection. The candidate will be involved with algorithm development, software design and implementation, and the development of experiments to test the algorithms.
**Prerequisites:**
Experience with Java/Android development, Planning and Tasking algorithms and Markov Decision Processes is a plus.

---

**Partial Face Recognition**
Faculty Advisor: TBD (Suggest William T. Freedman)
Mentor(s): Nicholas Orlans (MITRE)

table

<table>
<thead>
<tr>
<th>Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Area(s): Cross disciplines spanning Computer Science, Machine Learning, Computer Vision and Biometrics</td>
</tr>
<tr>
<td>Application Area: Biometric Identification Sciences</td>
</tr>
</tbody>
</table>

**Description of project and expected outcomes:**
Humans can recognize partial faces of familiar subjects [1], yet automated systems lack partial search capability. Various modeling and reconstruction techniques are used to estimate or hallucinate a complete face when one is not presented [2, 3, 4], yet reference applications and objective evaluation methods are not available.

Specifically, we seek to compare alternate methods for face fitting and reconstruction approaches as they apply to facial recognition challenges, determining their relative effectiveness as a prototype application. We also will assess the utility of preprocessing techniques for being able to submit partial face images to commercial solutions.

---

The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE’s concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.
The student will be expected to focus on the following tasks:
1. Review prior state of the art in face modeling.
2. Develop automated (or semi-automated) face fitting of partial images.
3. Demonstrate reference tools to process and display regional/partial recognition (e.g. recognition heat maps).
4. Propose evaluation protocols for partial face recognition.

[1] Example of two known partial faces we easily recognize and one unknown

![Example of two known partial faces we easily recognize and one unknown](image)


The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.


Prerequisites: Students should have an understanding of imaging, evaluating biometric systems, and programing in OpenCV/C++ and Python. Background work in machine learning and computer vision is desirable.

System Structure Morphology through Category Theory
Faculty Advisor: TBD
Mentor(s): Alan Stone (MITRE)
Contact:
Research Area(s): Systems Science, Mathematics, Computer Science

Research Area(s): Data-flow processing, graph theory, and category theory.

Application Area: Broadly – systems science. Focused demonstration - fault-repair synthesis in a distributed data-flow processing system.

Description of project and expected outcomes:

The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE’s concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.
Develop mathematical constructs based on Category Theory that allow one to define, compose, transform, and perturb (systematically and stochastically) data-flow processing graphs. Develop computational methods based on these constructs that allow one to check for equivalence, synthesize alternative equivalent graphs, and perform composition and reduction.

Once established, these structure morphism constructs may be used as the foundation for many common systems problems (e.g. adaptive function composition and distributed hardware mapping, fault tolerance, load balancing, and others). The proposed application of these constructs is to demonstrate service continuity in fault-fractionated systems through the use of repair synthesis. Briefly, as the data-flow graph is punctured by failure, one employs these new methods to synthesize an equivalent repair graph, that, when composed with the faulty one, results in equivalent overall function. This could be demonstrated on nearly any kind of distributed platform (e.g. a dozen Raspberry Pi’s or a server rack).

While its potential applications are interesting, the underlying goal is to explore the feasibility of moving from what are frequently graph-theoretical methods, to category-theoretical methods. If these methods prove successful and compelling, then the systems scientist will have a new tool for defining, analyzing, predicting, and synthesizing the structure of systems.

The student will be expected to focus on the following tasks:  
1. Survey current graph-theoretical methods for definition and
transformation of data-flow processing graphs.

2. Develop an understanding of category theory as it applies to its use for data-flow processing graphs.

3. Develop mathematical constructs that allow one to define, compose, transform, and perturb (systematically and stochastically) data-flow processing graphs.

4. Develop computational methods that allow one to check for equivalence, synthesize alternative equivalent graphs, and perform composition and reduction.

5. Develop a software demonstration that employs these methods to show how to perform synthesize fault repair in a distributed data-flow processing graph.

**Prerequisites:** Graph theory and category theory.

---

"Early Enterprise Attack Detection and Target Identification"

Faculty Advisor: TBD

Mentor(s): Dr. Alex Edward Raj (MITRE)

Contact:

Research Area(s): Cyber Security

**Description of project and expected outcomes:**

Large enterprises deploy many devices, such as routers, switches, Intrusion Prevention Systems (IPS), Intrusion Detection Systems (IDS), firewalls, Domain Name System (DNS), Web content filters, mail relays, web proxies and servers in the network, boundary and compute infrastructures in a secure layered architecture. The

---

The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.
compute infrastructures are well protected behind multiple boundaries along with client level security in a defense in depth architecture. When cyber-attacks are identified, operators identify the extent of the adversary’s progression in the enterprise infrastructure to effectively combat the adversaries without losing ground. Currently, operators spend a lot of time identifying the attack and locating the threat before remediating it. Automated early attack detection is the key for stopping adversaries before losing the assets.

Currently, researchers are looking at resources such as enterprise log, Simple Network Management Protocol (SNMP), and network statistics and Network Behavior Anomaly Detection (NBAD). These are being explored for early enterprise attack detection and location identification independently. These independent individual data sources cannot provide accurate, scalable and comprehensive early enterprise attack detection and target identification. An integrated, scalable, deployable algorithm and tool, that uses combination of these data are required for successfully detecting and locating attacks early in the large enterprise network.

**Prerequisites:** Motivated students with a strong interest in cyber security, algorithm development and some programming skills (Python, C or C++)

**The student on this task will be expected to:**
1) Learn basic cyber security threat detection methods (survey);
2) Get familiar with network, boundary and compute server device events and management information base (MIB) for the lead products in the market;
3) Learn how to locate cyber-attacks in a large enterprise;  
4) Develop threat detection algorithm;  
5) Develop threat target identification algorithm;  
6) Deploy these algorithms in the lab for demo;  
7) Identify product’s threat detection and identification requirements/deficiencies.

"Intermittent Mobile Web Services"  
Faculty Advisor: TBD  
Mentor(s): Patrick Maurer (MITRE)  
Contact:  
Research Area(s): Distributed Applications, Computer Science

**Description of project and expected outcomes:**

Historically, web services have been built with dedicated servers drawing upon a static set of resources, and using generic applications. Mobile web services often add a more specialized local application component. The emergence of cloud-based services and the “internet of things” (IoT) allows for new classes of distributed applications using dynamic processing and data resources. Enabling modern, cloud-based mobile web services over a time-varying network that includes mobile components (vehicles, pedestrians moving through a space, acting as both users and resources) creates challenges to the applications, resources and

The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.
underlying network. Applications and resources must be disruption-tolerant to withstand outages, and connections and networks may need to be application-aware. Network topology and application message-passing may need to be jointly optimized.

This research project will examine cloud service delivery in a dynamic mobile environment. Several areas of study are possible, ranging across modeling and simulation, optimization, protocols, and web services architectures and development.

**Prerequisites:** Motivated students with a strong interest in one or more of: distributed systems, internet of things, mobile web services, network modeling and optimization, and having some programming skills (e.g., Java, Python, C, C++, Matlab)

**The student on this task will be expected to:**
1) Learn basic distributed computing methods;
2) Survey commercial cloud computing services and architectures;
3) Learn basics of dynamic / mobile networks;
4) Model application behavior in mobile networks;
5) Develop and test performance of applications using distributed computing over mobile networks;

"Software Defined Underwater Acoustic Communication System"
Faculty Advisor: TBD
Mentor(s): Ballard Blair, Michael Wentz, Marcus Rahimpour

The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.
The underwater acoustic communication channel is extremely challenging due to a low speed of sound, wide spreading in both Doppler and delay, and extreme geographical variability. Given these challenges, research community adoption of acoustic communication technology has been slow.

Software defined radios allow for rapid development of radio frequency (RF) communication systems. Applying this same methodology to the underwater communication problem would expedite system development and open research and development to a wider community.

The goal of this project is to research a software defined environment to be run on embedded processors to enable real time underwater communication. Specifically, the student will explore use of open source software solutions like GNURadio and other software like Liquid SDR and Python libraries. The software shall have the capability for users to program in a variety of communication waveforms (BPSK, QPSK, MSK, etc.) and protocols. Integration of acoustic front-end and transducers onto processing hardware will enable transmission and reception capabilities. Further capabilities include modulation/demodulation schemes as well as timing synchronization. Lab demonstrations of successful communication will permit larger body of water trials.
Prerequisites: Motivated students with a strong background in embedded systems, acoustics, communication, electronics, and software (Python, C, or C++).

The student on this task will be expected to:
1) Generate a simulation software toolkit (in MATLAB or similar) for modeling an acoustic communication system and understanding the challenges and providing a baseline for development.
2) Software defined communication toolkit for acoustic communication. This will include the basic communication software with basic building blocks enabling incoherent/coherent acoustic communication
3) Demonstration of hardware communicating over a wire using developed hardware and software without acoustic frontend
4) Integration of acoustic front-end and transducers onto processing hardware
5) Add additional blocks to enable phase coherent communication
6) Demonstration of one-way communication through water

Blockchain For Enterprise Transaction Security (BETS)
Faculty Advisor: TBD
Mentor(s): Dr. Suresh Damodaran (MITRE)

Contact

Research Area(s): Secure Enterprise and hosted cloud applications

**ABSTRACT**

Blockchain is a permissionless distributed database used to support the bitcoin protocol. Blockchain offers the promise of tamperproof record keeping. A dedicated organization, R3CEV LLC, has been recently set up by financial companies to explore the use of blockchain technology for trading commercial papers. This work (1) will evaluate blockchain technology for securing transaction records in enterprise and hosted cloud applications, and (2) identify limitations or potential vulnerabilities in blockchain technology. The specific applications for blockchain we would like to consider are trustworthy computing through fraud prevention in generation and storage of IRS tax returns, electronic health records, and provenance records for intelligence and sensor data.

**PROBLEM**

Blockchain technology is being proposed for use in a variety of areas such as voter records to electronic ledgers; however, there is minimal work on possible applications in the government agencies, and vulnerabilities in this technology.

**RESEARCH OBJECTIVE**

Explore application of blockchain technology for securing enterprise and hosted cloud application records in the cloud. Identify any vulnerabilities in the use of blockchain for such purposes.

**TECHNICAL BACKGROUND**

"""I don’t have a strong opinion on cryptocurrencies, but I have a strong opinion on the blockchain as a solution for contracts and supply chains and the Internet of Things. I think bitcoin is an interesting application for blockchain but there are thousands of applications and wider use cases beyond that.""" – Jerry Cuomo, IBM Fellow.

Bitcoin protocol [1] has been around since 2009. Currently, a consortium of more than 25 banks, led by financial technology company R3CEV LLC, is developing a framework for applying blockchain technology for trading commercial papers [2]. Blockchain works like a huge, decentralized ledger for the digital currency bitcoin which records every transaction and stores this information on a global network so it cannot be tampered with. It’s this technology that banks feel can be utilized in areas from remittances to securities exchanges [3]. Barclays said in May that bitcoin
created a more "elegant solution than our current payment system" in many ways but it also "fell short" in a number of areas [3]. It is not clear in what ways it fell short. Open source projects exist for blockchain such as Openchain [4] and Open Ledger project [5] by IBM. Tamper proofing the records is important in healthcare record storage [6]. Indeed, Estonia has already applied blockchain technology for health care records [7]. Blockchain may also be applied for tamper-proofing tax returns [8], provenance records in Intelligence, and sensor data in DoD.

**PREREQUISITES**

Motivated students with a strong interest in cyber security, database (NoSQL), algorithm development and some programming skills (Java, JavaScript, or Python).

**TECHNICAL APPROACH**

1. Deploy open source implementations of blockchain such as OpenChain in MITRE.
2. Develop usecases for securing enterprise and cloud applications that need tamper proof storage of records such as electronic health records, IRS tax returns, provenance records of sensor and intelligence data.
3. Implement representative usecases on the blockchain code base, and explore vulnerabilities.

**IMPACT**

Blockchain is a new technology that promises tamperproof transaction records. Identifying vulnerabilities in this technology, and exploring its applications will allow MITRE to be ready when sponsors are ready to explore this technology. The blockchain technology has potential for application in Healthcare (patient records), IRS (tax returns), Elections (voting records), Intelligence (provenance data), and DoD (sensor data).

References:
[5] Open Ledger Project, IBM
[7] Estonian citizens will soon have the world’s most hack-proof health-care records, http://qz.com/628889/this-eastern-european-country-is-moving-its-health-records-to-the-blockchain/
The authors’ affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the authors.