

Sixteenth Annual
Student Research Forum

March 21 & 22, 2019

COMPUTER SCIENCE AND SCIENTIFIC COMPUTING

PRESENTATION SCHEDULE

Session Chair: *Dr. Yuanzhu Chen*

Head, Computer Science: *Dr. Minglun Gong*

Director, Scientific Computing: *Dr. Ron Haynes*

THURSDAY, MARCH 21, 2019, EN-2022

09:00 Opening Remarks

09:10 Promotech: A universal tool for promoter detection in bacterial DNA sequences,
Ruben Chevez

09:30 Time dimension in mobile data, **Ali Farrokhtala**

09:50 Changing ϵ -Greedy decay in Atari DQN, **Lucas Critch**

10:10 Using Machine Learning to Reduce Search Space in Large State and Action Spaces,
Rory Campbell

10:30 Characterization of transcription factors' dimers from ChIP-seq meta-data using Position
Weight Matrix Graphs, **Aida Ghayour**

BREAK (20 minutes)

11:10 Genome-wide prioritization of autism risk genes using network science and evolutionary
computing, **Yu Zhang**

11:30 A Recommendation system based on Social Temporal Metrics, **Shorouq Al-Eidi**

11:50 Vehicular Social Networks, **Elham Karimi**

12:10 Illuminating the search for Osteoarthritis predictors, **Kyle Nickerson**

12:30 Orbital Construction: A Reinforcement Learning Approach to Building Enclosures Using
Swarm Systems, **Caroline Strickland**

LUNCH (1 hour)

14:00 Multi-scale Convolution Aggregation and Stochastic Feature Reuse for DenseNets,
Mingjie Wang

14:20 Deep Reinforcement Learning for Real-time Strategy Game Artificial Intelligence, **Rick
Kelly**

14:40 A Comparative Study Of Feature Clustering On Gene Selection, **Hanieh Marvi**

15:00 **Closing Remarks**

FRIDAY, March 22, 2019, EN-2022

09:00 Opening Remarks

09:10 Flower Species Classification by Using Support Vector Machine Classifier, **Tian Wang**

09:30 Sports Modeling and Prediction using Evolutionary Algorithm, **Shengkai Geng**

09:50 Compression improves image classification accuracy, **Wilson Ozah**

10:10 A software tool for biomarker discovery workflow using metabolomic data, **Arshad Arafat**

10:30 Ground plane estimation based on Hidden Markov Model, **Chengsi Zhang**

BREAK (20 minutes)

11:10 sRNA Target Prediction, **Kratika Naskulwar**

11:30 Road Map to Safety & Health - Obesity to City Planning: An Automated Complex-Networks-Based Graph Reduction System to Measure Road Importance, **Ali Alfosoool**

11:50 Linear constraints Method to Solve Different Block Design Problems, **Rida Albira**

12:10 A study of Reduced Latin Squares and Orthogonal Latin Squares, **Omran Khalifa**

12:30 A monotonicity preserving multigrid method, **Dawei Wang**

LUNCH (1 hour)

14:00 Anomaly Payload Signature Generation System Based on Efficient Tokenization Methodology, **Alaa Alhowaide**

14:20 Automated Image-Based Quality Control of Thin-Film Molecularly Imprinted Polymers, **Andrew Way**

14:40 Hierarchically-fused Generative Adversarial Network for text to realistic image synthesis, **Xin Huang**

15:00 **Closing Remarks**

THURSDAY, MARCH 31, 2016

09:10 **Ruben Chevez**, M.Sc. Computer Science

Title: **Promotech: A universal tool for promoter detection in bacterial DNA sequences**

Abstract: A promoter is a genomic sequence where the transcription machinery binds to start copying a gene into an mRNA molecule. Detecting promoter sequences is important for microbiology since they play a central role in the regulation of gene expression. Most of the tools available focus on a single bacterium; or a set of sigma (σ) factors. Promotech is being developed to address this problem, offering a machine-learning-based classifier trained to generate a model that generalizes and detect promoters in a wide range of bacteria.

* * *

09:30 **Ali Farrokhtala**, Ph.D. Computer Science

Title: **Time dimension in mobile data**

Abstract: Many complex systems and modern technologies which provide a vast amount of data are providing us possibilities to explore, investigate and explain some challenges in various scientific areas, where all have one shared element, that is time. Temporal network analysis is an excellent approach to joint the time dimension of data with its present hidden patterns and presents insights about the behavior of the underlying system's components. We aim first to highlight some of the differences in the temporal network compared to static models. Furthermore, we show that not only the temporal patterns in the empirical systems are not random, but also a better understanding it, help us to find a different result from conventionally expected knowledge. One instance is the presence of a smaller range of node reachability among the people in a small community such as a university. Applying randomized null model, we also found that any small change in contact sequence distribution, including altering contact frequencies, redistributing them between users or destroying daily cycle of behavioral repeating patterns, significantly affect some of the network statistics, such as the network's average reachability ratio.

* * *

09:50 **Lucas Critch**, M.Sc. Computer Science

Title: **Changing ϵ -Greedy decay in Atari DQN**

Abstract: Deep reinforcement learning algorithms show promising results in learning

complicated policies in benchmark environments such as the Arcade Learning Environment (ALE). Deepmind has previously shown that their Deep Q-Network (DQN) is able to learn a variety of difficult tasks. Using pixels as input, their algorithm was capable of learning 49 Atari games at or above human level. This present work aims to replicate a subset of their findings to gain a better understanding of the topics at hand. Previous work has used epsilon values of 1 to 0.1 in their ϵ -greedy exploration of actions. We propose decaying the lower bound to 0.01 over the same period and comparing the results with those using the previous values. The implication is that a trained network should not continue to choose random actions after a given amount of time (10% to 1% random). This will help give a better understanding of DQN and the impact of greedy action selection, in hopes of further advances in deep reinforcement learning.

* * *

10:10 **Rory Campbell**, M.Sc. Computer Science

Title: **Using Machine Learning to Reduce Search Space in Large State and Action Spaces**

Abstract: Artificial intelligence (AI) is an integral part of both game development and play. AI that is dynamic enough to provide a level of challenge appropriate to a variety of players can improve a game's reception for both audiences and critics. Strategy games are often complex due to the large number of possible scenarios they present. Given how challenging strategy games are, the AI players in these games have been historically weak. Solutions to these complex problems have utility for AI in general, extending beyond the games in which they are implemented.

Machine Learning (ML) is a sub-field of AI which encompasses many analytical and learning techniques useful in writing a sophisticated AI game-player. As a testing platform for my work, I will be working with the Canadian video game company Lunarch Studios on incorporating new ML techniques into the existing AI framework of their retail video game, Prismata. By improving the AI system within Prismata, it will strengthen their product and provide their users with an overall better gaming experience. Due to the complex nature of the game, any technique capable of performing well in this environment should be applicable to other industries as well.

The existing Prismata AI is a searching agent as opposed to a learning agent. We are using supervised learning in an attempt to reduce the branching factor in the search. One of the central components in developing this protocol is constructing a state representation. Developing a concise but sufficiently comprehensive summary of a game state will greatly improve the efficiency of testing and the quality of learning. Once a state representation is achieved, we must carefully select an appropriate set of features and/or actions as the output for our learning. The work on lessening the search space can then begin.

* * *

10:30 **Aida Ghayour**, M.Sc. Computer Science

Title: **Characterization of transcription factors' dimers from ChIP-seq meta-data using Position Weight Matrix Graphs**

Abstract: Transcription factors (TFs) are key cellular components that control gene expression. They recognize specific DNA sequences, the TF binding sites (TF-BSs), and thus are targeted to specific regions of the genome where they can recruit transcriptional co-factors and/or chromatin regulators to fine-tune spatiotemporal gene regulation. Therefore, the identification of TFBSs in genomic sequences and their subsequent quantitative modeling is of crucial importance for understanding and predicting gene expression. The current golden standard technique for identifying TF's binding sites is ChIP-seq, which consist in performing a Chromatin Immunoprecipitation (ChIP) to pull-down all DNA fragment that interact with a protein of interest, following by high through-put DNA sequencing (seq). Using established motif finding programs, we can compute a DNA motif prototype recognized by a specific TF. Modeling these binding sites, are done with the help of Position Weight Matrices (PWM), a commonly used $4*n$ matrix that summarized the consensus binding sequence by keeping frequency of nucleotides at each position of sequence. These matrices are powerful tools for predicting the binding site of TFs in gene regulation studies. Although present method of TFBSs modelling is very convincing, it has some drawbacks. Most transcription factors do not work alone. Many large TF families form complex homotypic or heterotypic interactions through dimerization. For gene transcription to occur, a number of transcription factors must bind to DNA regulatory sequences. Because of this contribution of multiple TFs, a single sequence matrix/logo is not sufficiently expressing as it is not properly designed to illustrate the impact of other TFs engaged in transcription process. To tackle this issue, we purpose a model which consists a graph of multiple PWMs (or Seq-logos) to have a better representation of a binding site of a TF of interest, in the presence of other factors.

* * *

11:10 **Yu Zhang**, M.Sc. Computer Science

Title: **Genome-wide prioritization of autism risk genes using network science and evolutionary computing**

Abstract: Autism spectrum disorder (ASD) is a neuropsychiatric disorder characterized by impairments in reciprocal social interaction and communication, and the presence of restricted and repetitive behaviors. ASD is predominantly heritable, but the underlying genetic determinants are still largely unknown. Although over one thousand genes have been estimated to shown association with ASD, only a small fraction are known with strong genetic evidence from sequencing studies. In addition, the affected neural circuits and cell types of ASD individuals remain unclear and may

vary at different developmental stages. Over the past decades, rapid advances in network biology indicate that molecular networks offer a powerful conceptual framework that could potentially revolutionize our view of disease pathologies. Moreover, genetic programming (GP), an evolutionary algorithm, is an advanced and intelligent analytical tool, which can be used to unravel the disease-gene association with great potential. In our study, we propose to predict ASD-associated genes on a genome-wide scale using network science and evolutionary computing. Following the hypothesis that genes similar to known ASD-associated genes may potentially contribute to ASD susceptibility as well, we use biological networks of physical and functional gene interactions to characterize structural properties of genes as nodes in the networks. Then, by exploiting the structural properties, we use GP to learn predictive models that can prioritize novel candidate genes to what extent they are involved in ASD manifestation and help facilitate the downstream biological research to refine our understanding of the aetiology of ASD, in order to improve clinical diagnosis, treat and even intervention of the disease

* * *

11:30 **Shorouq Al-Eidi**, Ph.D. Computer Science

Title: **A Recommendation system based on Social Temporal Metrics**

Abstract: Recommendation systems have changed the way people find products, locations of interest, services or new friends. The technology behind recommendation systems has evolved to provide users preference and social influences. In my work, I present an approach to developing a recommendation system based on the user similarity and temporal metrics applied to graphs that represent the objects characteristics, user profiles and influences obtained from social connections. It exploits temporal graph centrality measures to elaborate personalized recommendations from the semantic knowledge represented in the graph. The time-ordered bipartite graph model and selected graph algorithms for calculating graph centralities that are the core of the recommender system are presented. Semantic concepts such as semantic predominance and similarity measures are adapted to the graph model.

Keywords: recommender systems, social metrics, graph models.

* * *

11:50 **Elham Karimi**, Ph.D. Computer Science

Title: **Vehicular Social Networks**

Abstract: Vehicular Social Networks (VSNs) are social networks shaped on roads by drivers or passengers who have social relationships, communications, and common interests. The utilisation of vehicular users' social properties brings well networking and social provision to novel applications and services. This work overviews recent

achievements in VSN. Some significant social metrics such as closeness and betweenness play a crucial role to confirm if vehicular network presents a social manner. We usage real dataset to describe the mobility of a set of vehicles. Finally, we consider new social points of view in vehicular networks.

Key words: vehicular networks, social network analysis, dynamic sociality

* * *

12:10 **Kyle Nickerson**, M.Sc. Computer Science

Title: **Illuminating the search for Osteoarthritis predictors**

Abstract: Machine learning is a powerful technique for discovering and modeling relationships between predictive features and outcomes, by rapidly searching a space of possible models and finding one which bests fits the available training data. Recently some researchers have begun to argue that instead of focusing on finding a single best performing model for a dataset, machine learning should take a more holistic view, and begin to place more emphasis on discovering how high performing solutions are distributed throughout the search space. These more holistic machine learning algorithms are known as illumination algorithms- and in contrast to traditional search algorithms- aim to provide a diverse set of simple solutions, rather than a single complex solution which performed best on a training dataset.

The benefits of our approach are twofold. First, predictions from the collection are better able to generalize to new data not used to train the model. Second, the collection provides a set of simple relationships between features and predictions, which can easily be interpreted by humans. This increased interpretability can help researchers place more trust in the model predictions, as well as provide indications if a model is making biased- or otherwise inappropriate predictions.

We applied our technique to predict the presence of osteoarthritis (OA), using patient blood test data. Results show our method was able to find much simpler models than a previous study using this data. In addition, we show these simpler models are more accurate at predicting outcomes of patients not used to train the model.

* * *

12:30 **Caroline Strickland**, M.Sc. Computer Science

Title: **Orbital Construction: A Reinforcement Learning Approach to Building Enclosures Using Swarm Systems**

Abstract: Orbital construction - an algorithm developed by Dr. Andrew Vardy coordinates swarms of simple robots (separated into two distinct groups) to build enclosures

using a projected scalar field. Inspired by the use of template pheromones in insect collective construction, this algorithm takes advantage of circumnavigation, periphery seeking, and nudging objects inwards to achieve a projected shape. By applying a relatively straight-forward control algorithm to a large number of simple agents, the end result is often fairly complex.

Reinforcement learning is a subfield of machine learning that allows one or more agents to learn the best action to take from experiencing and interacting with their environment. Such systems are extremely well-suited to robotics, as agents often interact with complex environments through a variety of unique sensors, actuators, and techniques. When reinforcement learning is applied to a multi-agent system, it is called 'multi-agent reinforcement learning' (MARL). A main feature that MARL offers is flexibility - a multi-agent decentralised system can have agents added, removed, or reconstructed without need for rewriting the system. This allows for more robust solutions due to its ability to cope with failure.

With the use of simulators paired with MARL, we can effectively learn policies that result in optimal actions for pattern formation. This would be a vast improvement, as it removes dependence on hard-coded actions and allows us to define our own state representation. As well, reinforcement learning eliminates the need for writing control algorithms in the first place - which tend to be extremely task-specific and time-consuming. Learning a policy takes into account a variety of variables that must be predetermined, which grants the user control of specific aspects before learning begins.

* * *

14:00 **Mingjie Wang**, Ph.D. Computer Science

Title: **Multi-scale Convolution Aggregation and Stochastic Feature Reuse for DenseNets**

Abstract: Recently, Convolution Neural Networks (CNNs) obtained huge success in numerous vision tasks. In particular, DenseNets have demonstrated that feature reuse via dense skip connections can effectively alleviate the difficulty of training very deep networks and that reusing features generated by the initial layers in all subsequent layers has strong impact on performance. To feed even richer information into the network, a novel adaptive Multi-scale Convolution Aggregation module is presented in this paper. Composed of layers for multi-scale convolutions, trainable cross-scale aggregation, maxout, and concatenation, this module is highly non-linear and can boost the accuracy of DenseNet while using much fewer parameters. In addition, due to high model complexity, the network with extremely dense feature reuse is prone to overfitting. To address this problem, a regularization method named Stochastic Feature Reuse is also presented. Through randomly dropping a set of feature maps to be reused for each mini-batch during the training phase, this regularization method

reduces training costs and prevents co-adaptation. Experimental results on CIFAR-10, CIFAR-100 and SVHN benchmarks demonstrated the effectiveness of the proposed methods.

* * *

14:20 **Rick Kelly**, M.Sc. Computer Science

Title: **Deep Reinforcement Learning for Real-time Strategy Game Artificial Intelligence**

Abstract: Games have long been considered important Artificial Intelligence (AI) testbeds, from Chess to Go and now to more complex video games such as the real-time strategy (RTS) game StarCraft II. RTS games are military simulations played on 2D maps with up to hundreds of units to control and include hidden information or “fog of war”. The real-time nature of RTS games and incomplete information is closer to real-world situations than classic board games, making them good environments for AI advancement. The current state-of-the-art in RTS AI involves hand-crafted strategies using expert human knowledge mixed with various AI techniques such as heuristic search. These techniques are not easily transferred to other domains, and skilled human players can defeat these AI players by adapting to their predictable strategies. Reinforcement learning (RL), an AI method in which a system learns to maximize a reward by interacting with its environment, has in recent years been combined with deep neural networks and applied to games with large state spaces, such as Go. The StarCraft II Learning Environment has recently been released, enabling deep learning in that game environment. Our goal is to use RL methods to train an agent in micromanagement combat tasks in StarCraft II.

* * *

14:40 **Hanieh Mavi**, M.Sc. Computer Science

Title: **A Comparative Study Of Feature Clustering On Gene Selection**

Abstract: Cancer/disease datasets normally contain a small size of samples versus a large number of genes as features. Selecting relevant genes which are involved with the different type of diseases is a challenging task. As a means of dimensionality reduction, feature selection is used for selecting a subset of features which leads to better performance of the classifier. However, most feature selection algorithms suffer from high computational complexity. To tackle this issue, in this project, we intend to incorporate feature clustering in gene selection. First, genes are clustered into gene groups with a clustering algorithm so that genes belonging to the same cluster have similar expression profile. Then, from each cluster, the representative gene(s) are selected. After selecting representative genes from all clusters, feature selection is applied to this set of representatives. This method decreases the

computational complexity and redundancy of genes which leads to better classification accuracy and finding informative genes. In this research, we conduct a comprehensive study by examining widely used clustering, feature selection, and classification methods, and applying them to different cancer/disease datasets. Results in this study will show that with incorporating feature clustering to gene selection, computational complexity will be reduced significantly and better classification accuracy will be achieved.

* * *

15:00

Closing Remarks

FRIDAY, MARCH 22, 2019

09:00 Opening Remarks

09:10 **Tian Wang**, M.Sc. Computer Science

Title: **Flower Species Classification by Using Support Vector Machine Classifier**

Abstract: In the last few years, researchers in both computer science and biological fields have built different models to recognise flower species. Due to the similar properties such as colour, shape and structure, it is crucial to discriminate flowers effectively. Using support vector machine classifier with multiple kernel framework is an excellent strategy to get better performance with considering the features selection. Depends on the past experiments, Flower17 and Flower102 datasets will be used as the elementary image library. Based on the features of the flower, we split the color model as foreground (petal, sepal, and tepal)and background (greenery) . There are a bunch of models we will use to achieve the selection performance. Deep convolutional neural networks, the minimum redundancy maximum relevance method and radial bases function kernel are all resulted in high accuracy (over 90%) even after input the same classes.

* * *

09:30 **Shengkai Geng**, M.Sc. Computer Science

Title: **Sports Modeling and Prediction using Evolutionary Algorithm**

Abstract: NBA Playoffs are largely enjoyed by fans around the globe. In this project, I propose to use an evolutionary algorithm, namely genetic programming, to model and predict the final result of the NBA Playoffs. I aim at using the regular season performance statistics of each team to predict their final ranks in the Playoffs. Historical data of NBA teams were collected in order to train the predictive models using genetic programming. The preliminary results show that the algorithm is able to achieve good prediction accuracy, as well as to provide importance assessment of various performance statistics in determining the probability of winning the final championship.

* * *

09:50 **Wilson Ozah**, Ph.D. Computer Science

Title: **Compression improves image classification accuracy**

Abstract: We study the relationship between the accuracy of image classification and the level of image compression. Specifically, we look at how various levels of JPEG and SVD

compression for classical images from image compression literature affect the score of the correct answer in Inception-v3, a TensorFlow-based image classifier trained on the ImageNet database. Surprisingly, the compression seems to improve the ability of Inception-v3 to recognize images, with the best performance seen at fairly high degrees of compression for most images tested (with two-thirds achieving maximal score at JPEG quality under 20, corresponding to more than tenfold reduction in file size). The same behaviour holds for images compressed using the singular value decomposition (SVD) method. This phenomenon suggests that even significant compression can be beneficial rather than detrimental to image classification accuracy, in particular for convolutional neural networks. Understanding when and why compression helps, and which compression algorithm and compression ratio are optimal for any given image remains an open problem.

* * *

10:10 **Arshad Arafat**, M.Sc. Computer Science

Title: **A software tool for biomarker discovery workflow using metabolomic data**

Abstract: Metabolomics is the quantitative study of small-molecule metabolites and their underlying biochemical processes. Comparing to the more developed omics family such as genomics and proteomics, metabolomics most directly reflects cells' phenotypes. This newly popular domain can be used in various areas starting from biomedical research to interaction studies between organisms and their environment.

One of the tasks of metabolomics is biomarker discovery for human diseases. However, this domain still lacks integrative and advanced tools and techniques for practical analyses and applications. One example is the absence of a complete software tool focusing on data analysis for biomarkers discovery.

By reviewing various existing and related software tools we found that the research community can really benefit a software tool that focuses on metabolic marker discovery and is able to provide a complete workflow for data preprocessing and metabolic marker selection. In this project, we propose to develop a web-based software tool for metabolic biomarker discovery that can be used by a large research community.

* * *

10:30 **Chengsi Zhang**, M.Sc. Engineering

Title: **Ground plane estimation based on Hidden Markov Model**

Abstract: Constructing 3D environment from calibrated multiple 3D sensors is very time consuming with the current approaches. I believe that finding the ground plane, and

then estimating the orientation and relative position of each camera should alleviate the limitations of multi-sensor calibration. In order to achieve the ground plane estimation with minimum preconditions, such as sensors orientation or sensors movement, I formulate the problem as a state-continuous Hidden Markov Model (HMM) where the hidden state contains $\rightarrow t$ and $\rightarrow n$ that are estimated by sampling and decomposing homographies based on 2D video data, which then will be extended to 3D point cloud data.

* * *

11:10 **Kratika Naskulwar**, M.Sc. Computer Science

Title: **sRNA Target Prediction**

Abstract: Bacterial small regulatory RNAs (sRNAs) play a vital role in the regulation of gene expression in bacteria. sRNAs influence gene expression by interacting with mRNAs or proteins. Bacterial sRNAs are involved in a variety of bacterial functions like environmental stress response, metabolism and virulence. In order to understand the functional roles of sRNAs, we need to identify the mRNAs and/or proteins that these sRNAs interact with. These mRNAs or proteins are called targets of the sRNAs. There are several computational tools available for sRNA target prediction, however, these tools have a high number of false positives and hence there is room to increase their accuracy. Our research centres on building up a machine learning based method for sRNA target prediction that improves over existing tools. As a result of our research, we expect to develop a more accurate method for sRNA target prediction applicable to any bacterium that will be available for microbiologists.

* * *

11:30 **Ali Alfosool**, Ph.D. Computer Science

Title: **Road Map to Safety & Health - Obesity to City Planning: An Automated Complex-Networks-Based Graph Reduction System to Measure Road Importance**

Abstract: Road networks are abundantly available from various sources such as OpenStreetMap and are great recourses for various types of analysis such as association between neighborhoods' walkability and local obesity rate, or road characteristics and city planning. Unfortunately, existing works focus on single field requiring rigorous manual procedures. This research aims to develop and implement algorithms to automate data processing, graph reduction while preserving network's original structure, application of different centrality measures, and applying non-network characteristics such as amenities, elevation, type, condition and structure of roads and their neighborhood, etc., to compute scores for each road in the network. Our goal is to generalize the automated system in such a way it automatically

processes and reduces the graph, then given a set of road characteristics, produces customizable scores appropriate to a wide range of applications. Such as walkability scores to promote healthier lifestyles by encouraging walking or help with city planning to develop better more walkable city structures and to help our environment by reducing vehicular use. We have been able to achieve up to 70% node reduction on a city-level data set (i.e. Toronto). Our algorithms are tested and verified by numerous artificial graphs, special cases and real road networks. Some preliminary walkability scores are available for Downtown St. John's and Downtown Toronto.

* * *

11:50 **Rida Albira**, Ph.D. Computer Science

Title: **Linear constraints Method to Solve Different Block Design Problems**

Abstract: Combinatorial design theory studies finite sets satisfying specific combinatorial properties. In particular, any design is a pair (X, A) , such that X is a set of elements called points, and A is a multiset of non-empty subset of X called blocks. They have numerous uses from experimental designs to cryptography, geometry and deep learning. In spite of extensive work in this field, there are still open questions regarding existence of designs for some sets of parameters. In this work we look at a heuristics for determining existence of balanced incomplete block designs (BIBD), as well as packing, covering and nuclear design. In particular heuristics used in satisfiability (SAT) solving. Previously, many methods have been implemented include backtracking Gibbons 1976, hill climbing Stinson 1985 to solve some block design problems (BIBD in particular). The early work encoding these problems as instances of satisfiability problem was due to Prestwich 2001, and several related problems were solved by SAT techniques. The result of comparing different algorithms such as (hybrid algorithm of neural network and simulated annealing NN-SA by Bofill et al. 2003, local search LS algorithm by Prestwich 2001, branch and bound BB, Tabu search TS algorithm by Yokoya and Yamada 2011, and linear integer programming LIP method which use SAT solver Mandal 2015) shows that the LIP algorithm solved highest number of cases from all other listed algorithms. So far, we have showed that, even with the recent improvements in SAT solving, using Prestwich encoding quickly becomes infeasible. we are recently investigating other types of encoding and solver heuristics, in particular pseudo-Boolean encoding. In this recent work, the focus will be as mentioned earlier on using pseudo-Boolean constraints to solve the four types of block designs. The balanced incomplete block design (BIBD), packing, covering and nuclear designs.

* * *

12:10 **Omran Khalifa**, M.Sc. Computer Science

Title: **A study of Reduced Latin Squares and Orthogonal Latin Squares**

Abstract: A Latin Square (LS) of side n can be described as an $n \times n$ array in which n distinct objects are arranged so that each object occurs exactly once in each row and each column. LS is reduced if the first row and the first column are in standard order. Two Latin Squares L_1 and L_2 of side n , said to be orthogonal if they are superimposed, so the resulting set of n^2 ordered pairs are distinct. This is an example of a more general problem in combinatorics, in that is, given a set of objects belonging to a finite set to find arrangements of those objects in accordance with certain restrictions. Latin Squares are used in various fields such as design theory, experimental design, coding, and cryptography. Working on LS is an interesting task, while at the same time constructing Latin Squares of the same side and finding the relationships between them is a challenge. Moreover, so far we do not know the number of reduced Latin Squares of side greater than 11 ($n > 11$). The reason may be because are dealing with extremely large numbers, even for reasonably small values of n .

In this research, our goal is to study existing Latin Squares in more depth to figure out their properties. We use new techniques to break down the structure of reduced Latin Squares, along with combinatorial analysis in order to write an algorithm to find the number of reduced Latin Squares of side greater than 11, ($n > 11$), then a hill-climbing algorithm to construct some sets of them. Results in this study will show, a different way of solving such this problem.

* * *

12:30 **Dawei Wang**, M.Sc. Scientific Computing

Title: **A monotonicity preserving multigrid method**

Abstract: In this presentation, I will show you a multigrid method that is able to preserve the monotonicity of the exact solution for 1D diffusion equation. This research is motivated by the nontangling requirement in adaptive mesh generation.

* * *

14:00 **Alaa Alhowaide**, Ph.D. Computer Science

Title: **Anomaly Payload Signature Generation System Based on Efficient Tokenization Methodology**

Abstract: Signature-based intrusion detection systems are widely used as an efficient network security control. Unfortunately, security experts manually craft attack signatures after capturing and analyzing the exploit code. Therefore, those systems are only able to detect known attacks. In this paper, we propose anew automated and content-based signature generation system that generates anomaly profiles to detect new and previously unknown attacks and worms. The proposed system, denoted SCANS uses a natural tokenization method that speeds up the signature generation process by

producing a fewer number of substrings. In this system, we propose a new stop character technique that will help to overcome signatures' substrings granularity limitations of the old stop word techniques. In addition, SCANS introduces an improved normalized binary detection model specifically tailored for attacks detection. Experimental testing using DARPA IDS dataset shows a 95% malicious packets detection rate for port 23, with specificity of 88.4% and 94.6% for ports 21 and 25, respectively

* * *

14:20 **Andrew Way**, M.Sc. Scientific Computing

Title: **Automated Image-Based Quality Control of Thin-Film Molecularly Imprinted Polymers**

Abstract: We present results of applying a feature extraction process to images of thin-film Molecularly Imprinted Polymers (MIPs) on glass substrates for defect detection. Features such as MIP side lengths, aspect ratio, internal angles, and edge roughness are obtained by using Hough transforms, Canny edge detection, and intensity images. A Self Organizing Map (SOM) is used for classification of texture of MIP surfaces. The SOM is trained on a data set comprised of images of manufactured MIPs. The raw images are first processed using Hough transforms and Canny edge detection to extract just the MIP-coated portion of the surface, allowing for surface area estimation and reduction of training set size. The training data set is comprised of 20-dimensional feature vectors, each of which is calculated from a single section of a gray scale image of a MIP. Haralick textures are among the quantifiers used as feature vector components. The training data is then processed using principal component analysis to reduce the number of dimensions of the data set. After training, the SOM is capable of classifying texture, including defects. These preliminary results for texture analysis via the SOM and measurement of geometric properties via Hough transforms and Canny edge detection are encouraging.

* * *

14:40 **Xin Huang**, Ph.D. Computer Science

Title: **Hierarchically-fused Generative Adversarial Network for text to realistic image synthesis**

Abstract: Generating photorealistic images from text descriptions is a challenging problem that has many applications in computer vision. While existing approaches on this topic have achieved impressive success, to generate 256x256 images from captions, they commonly resort to coarse-to-fine scheme and associate multiple discriminators in different stages of the networks. Such a strategy is both inefficient and prone to artifacts. In this proposal, we present an end-to-end network that can generate 256 x

256 photo-realistic images with only one discriminator. We fully exploit the hierarchical information from different layers and directly generate the fine-scale images by adaptively fusing features from multi-hierarchical layers. We quantitatively evaluate the synthesized images with Inception Score, Visual-semantic Similarity and average training time on the CUB birds, Oxford-102 flowers, and COCO datasets. The results show that our model is more efficient and noticeably outperforms the previous state-of-the art methods.

Keywords: Generative Adversarial Networks; text-to-image synthesis; hierarchical features fusion

* * *

15:00

Closing Remarks