SDAS Previous Seminars 2017

Patrick Emami, (PhD student - Computer Science) *Title:* An Introduction to Deep Learning *Abstract:* The deluge of data brought on by the assimilation of technology into all aspects of society has resulted in corresponding advances in our abilities to process, analyze, and extract patterns from data. Over the past 5-10 years, Deep Neural Networks have cemented their position as the state-of-the-art approach for many statistical learning tasks. In this presentation, I will provide an overview of 1) the key breakthroughs that enabled the successes of Deep Learning and 2) the technical concepts underlying modern Deep Learning systems. Finally, I will present some examples showcasing the diverse set of problems that Deep Learning can be applied to. September 7, 2017

NikolaMilicevic (PhD student - Mathematics) *Title:* Persistent Homology of Subsamples September 14, 2017

Leo Betthauser (PhD student - Mathematics) *Title*: Digital Imaging from a Topological Perspective *Abstract*: Topological Data Analysis (TDA) is a mathematical field that utilizes qualitative features of shapes in order to distinguish high dimensional data sets. Recently TDA has demonstrated potential to aid in shape reconstruction. Unfortunately, these reconstructions can be computationally taxing. I will present my ongoing research which simplifies these reconstructions using the discrete structure of voxel data to offer a computationally efficient solution. September 21, 2017

Peyman Jalali (PhD student - Statistics) *Title:* An Introduction to Graphical Lasso September 28, 2017

Mohammad Kazem Shirani Faradonbeh (Postdoctoral Research Associate Statistics) *Title:* Finite Sample Estimation in Non-stationary Vector Autoregressive Models *Abstract:* Estimation of the parameters for stationary Vector Autoregressive (VAR) models is a well-studied problem, both in the low and high-dimensional settings. However, there are hardly any results for the unstable case, especially regarding finite sample bounds. For this setting, classical results on least-squares estimation of the VAR parameters are not applicable and therefore new concepts and technical approaches need to be developed to address the issue.
Unstable VAR models reflect key real applications in engineering, econometrics, and finance. This study establishes finite sample bounds for the estimation error of the least-squares VAR estimates for a fairly large class of heavy-tailed noise distributions, and transition matrices. The results relate the sample size required as a function of the problem dimension and key characteristics of the true underlying transition matrix and the noise distribution. To obtain them, appropriate concentration inequalities for random matrices and for sequences of martingale differences are leveraged. November 2, 2017

Tamzidul Hoque (Ph.D. Candidate Computer Science, FICS) *Title:* A Systematic Feature Selection Methodology for Machine Learning Based Hardware Trojan Detection *Abstract:* Design houses often integrate Intellectual Property (IP) cores obtained from third-party vendors to reduce hardware design costs. While the design could be verified for a specified functionality, it is extremely hard to guarantee that no hidden, and possibly malicious capability exists in form of a hardware Trojan in the untrusted third-party IP (3PIP) blocks. While Trojan insertion at the foundry could be tackled to a certain extent due to the presence of a golden design, detection of malicious functionalities – i.e. the trust verification in 3PIP is a more intricate challenge since often nothing but the specification of the intended design is available. Several countermeasures have been proposed earlier most of which identifies a group of suspect nets or gates based on certain functional or structural properties that are commonly observed in publicly available hardware Trojan examples. While various machine learning classifiers could be trained to diagnose suspect 3PIPs for detecting Trojans based on such properties, ad-hoc selection of Trojan properties would impact the detection capability of the classifier. Besides, the presence of redundant properties increases the runtime of the classifier without adding any value. In this work, for the first time, we introduce a systematic methodology to select among various functional and structural Trojan properties for Trojan detection in 3PIP. We implement two different machine learning-based feature selection methods and observe the detection capability of the naive Bayes classifiers under the selected properties. By choosing the features systematically, false positive reduction of around 53% is achieved compared to the worst random selection of equal number of properties. This technique allows the

separation of properties based on Trojan models to further improve the detection capability and presents the correct property selection strategy based on the computational resources available. November 2, 2017

Shuhang Chen (Mathematics) *Title:* Influence Prediction for Continuous-Time Information Propagation on Networks *Abstract:* This talk presents a model predicting the time evolution of influence, defined by the expected number of activated (infected) nodes, given a set of initially activated nodes on a propagation network. To address the significant computational challenges of this problem on large heterogeneous networks, we establish a system of differential equations governing the dynamics of probability mass functions on the state graph where each node lumps a number of activation states of the network, which can be considered as an analogue to the Fokker-Planck equation in continuous space. We provides several methods to estimate the system parameters which depend on the identities of the initially active nodes, network topology, and activation rates etc. The influence is then estimated by the solution of such a system of differential equations. Dependency of prediction error on parameter estimation is established. This approach gives rise to a class of novel and scalable algorithms that work effectively for large-scale and dense networks. Numerical results are provided to show the very promising performance in terms of prediction accuracy and computational efficiency of this approach. November 16, 2017

Alex Wagner (Ph.D. Candidate Mathematics) *Title:* TBA November 30, 2017