samsi
NSF·Duke·NCSU·UNC

GDRR Program Opening Workshop
August 5-9, 2019

SPEAKER TITLES/ABSTRACTS

Melike Baykal-Gursoy
Rutgers, State University of New Jersey

"Balancing Performance and Security in Interference Assisted Secret Communication Networks"

Cyber-attacks are becoming major problems that involve individuals, organizations and even nations in the 21st century. In this paper, we consider a communication network in which individual parties are engaged in secret communication on multiple parallel channels. A defender, by applying interference to a potential eavesdropper, can prevent the eavesdropper to have access to confidential information on any of the channels. We utilize game theory to investigate the optimal strategies in this interference assisted secret communication network between a network defender and an eavesdropper. Two nonzero sum Nash games against two different types of attackers are studied. The first type of attacker tries to maximize his expected eavesdropping capacity and the other type of attacker tries to avoid being interfered by the defender. We derive detailed conditions for the existence of a unique Nash equilibrium for both games and present both equilibria and value functions explicitly. We then consider a Bayesian game in which the defender has incomplete information about the attackers' type. In all three games, our analyses show that there exists a cut-off index determining the set of communication channels that will be covered in the equilibrium strategies. Furthermore, the adverse effect of the interference signal plays an important role in the value of the cut-off index and it represents the vulnerability of wireless communication networks against eavesdroppers.

Andrew Belmonte
Huck Institutes of the Life Sciences, Pennsylvania State University

"Evolutionary Games and Adversarial Systems - Fundamentals and New Directions"

This talk will present an overview and summary of some of the fundamental aspects of evolutionary game theory, focusing on mathematical approaches taken to modeling mutually competitive interactions using differential equations and agent-based models. We will discuss social dilemmas that can occur when individual, selfish motivations conflict with the common or public good, as well as developments to meet these challenges. We will also address the role of individual connectedness on strategy evolution, ranging from random associations in well mixed populations to specified interactions in fixed or evolving networks. Along the way, we will highlight differences and similarities between decision makers and game players in these contexts.

Eunshin Byon
University of Michigan

“Variance Reduction for Reliability Assessment with Stochastic Computer Models”

Importance sampling has been widely used to improve the efficiency of deterministic computer simulations where the simulation output is uniquely determined, given a fixed input. To represent complex system behavior more realistically, however, stochastic computer models are gaining popularity. Unlike deterministic computer simulations, stochastic simulations produce different outputs even at the same input. This extra degree of stochasticity presents a challenge for reliability assessment in engineering system designs. Our study tackles this challenge by providing a computationally efficient method to estimate a system's reliability. Specifically, we derive the optimal importance sampling density and allocation procedure that minimize the variance of a reliability estimator. The application of our method to a computationally intensive, aeroelastic wind turbine simulator demonstrates the benefits of the proposed approaches.

Alicia Carriquiry
Iowa State University

“Risk and Decisions in the Criminal Justice System”

"Every criminal leaves a trace". This is known as the Lockhart Principle, and is the raison d'etre for crime scene investigators and forensic scientists, tasked with linking -- or not -- a suspect to a crime scene. Decision makers in Court are jurors, who are asked to choose between two alternative propositions: the defendant is, or is not, the perpetrator of the crime. To arrive to this decision, jurors must combine their own prior beliefs about the suspect's culpability with the information provided typically piece-meal by expert witnesses and others offering testimony. Experts, including forensic scientists, in turn face a sequence of decision nodes with uneven degrees of information and uncertainty to inform those decisions. Yet, decisions by experts and juries have tremendous impact on the fair administration of justice. We discuss the sequence of decision nodes and the gaps in knowledge that magnify the risks associated with those decisions, in the context of pattern evidence such as shoeprints and ballistics.

Daniel Eisenberg
Naval Postgraduate School

“Rethinking Resilience Analytics”

The concept of â€œresilience analyticsâ€ has recently been proposed as a means to leverage the promise of big data to improve the resilience of interdependent critical infrastructure systems and the communities supported by them. Given recent advances in machine learning and other data-driven analytic techniques, as well as the prevalence of high-profile natural and man-made disasters, the temptation to pursue resilience analytics without question is almost overwhelming. Indeed, we find big data analytics capable to support resilience to rare, situational surprises captured in analytic models. Nonetheless, this article examines the efficacy of resilience analytics by answering a single motivating question: Can big data analytics help cyberâ€“physicalâ€“social (CPS) systems adapt to surprise? This article explains the limitations of resilience analytics when critical infrastructure systems are challenged by fundamental surprises never conceived during model development. In these cases, adoption of resilience analytics may prove either useless for decision support or harmful by increasing dangers during unprecedented events. We demonstrate

that these dangers are not limited to a single CPS context by highlighting the limits of analytic models during hurricanes, dam failures, blackouts, and stock market crashes. We conclude that resilience analytics alone are not able to adapt to the very events that motivate their use and may, ironically, make CPS systems more vulnerable. We present avenues for future research to address this deficiency, with emphasis on improvisation to adapt CPS systems to fundamental surprise.

Tahir Ekin
Texas State University

"Fraud Analytics"

Fraud instances are seen in a wide range of domains such as finance, telecommunications and health care. In addition to the monetary loss, fraud results in loss of confidence to the governmental systems and diminishes the overall system quality. For instance, in health care, while overpayments are estimated to correspond up to ten percent of expenditures, they also have direct adverse impacts on patient health. Statistical methods have become crucial to handle overpayments especially given the increasing size and complexity. This talk aims to provide an overview of the challenges and opportunities of using analytical methods for fraud assessment, with an emphasis on health care systems. First, various fraud data types will be illustrated with some real world examples. Next, sampling, overpayment estimation and data mining methods will be discussed. In particular, the use of data mining methods have gained momentum to reveal hidden relationships and fraud patterns as well as for classification and prediction. The talk will conclude with a discussion of the proposed extensions in relevant domains such as handling improper payments and potential future work to address the practical needs such as integrated decision making.

Sujit Ghosh
North Carolina State University

"When are PH, AFT and PO Models not Adequate for Health Risk Assessment?"

In many clinical applications of health risk assessments using survival analysis, the commonly used semiparametric models, e.g., proportional hazards (PH), proportional odds (PO), accelerated failure time (AFT) etc. may turn out to be stringent and unrealistic, particularly when there is scientific background to believe that survival curves under different covariate combinations will cross during the study period. This talk presents an overview of various classes of nonparametric regression model for the conditional hazard function. In particular, a relatively new methodology is presented that has three key features: (i) the smooth estimator of the conditional hazard rate is shown to be a unique solution of a strictly convex optimization problem for a wide range of applications; making it computationally attractive, (ii) the model is shown to encompass a proportional hazards structure, and (iii) large sample properties including consistency and convergence rates are established under a set of mild regularity conditions. Empirical results based on several simulated data scenarios indicate that the proposed model has reasonably robust performance compared to other semiparametric models particularly when such semiparametric modeling assumptions are violated.

Evans Gouno
Universit´e de Bretagne Sud

"Bayesian Inference for Common Cause Failure Rate Based on Causal Inference with Missing Data"

A common-cause failure (CCF) is defined as the simultaneous failure of two or more components of a system due to a shared cause. The number of components involved in the system failure is called the order of the CCF. The cause of a CCF can be of different natures. For example, extreme environment or human error or manufacturing error can provoke synchronized failure.

In this talk we will describe a methodology to handle the causality to make inference on common-cause failure in a situation of missing data. The data are collected in the form of contingency table but the available information are only the numbers of CCF of different orders and the numbers of failure due to a given cause. Therefore only the margins of the contingency table are observed; the frequencies in each cell are unknown. Assuming a Poisson model for the count, we suggest a Bayesian approach and we use the inverse Bayes formula (IBF) combined with a Metropolis-Hastings algorithm to make inference on the rate of occurrence for the different combination cause, order. The performance of the resulting algorithm is evaluated through simulations. A comparison is made with results obtained from the _-composition approach to deal with causality suggested by Zheng et al. (2013).

Keyword: Common-cause failure, Inverse Bayesian formula, Contingency table, Missing data, Causal inference.

Joe Halpern
Cornell University

"Computer Science Meets Game Theory: Implementing Mediators Robustly"

The question of whether a problem in a multiagent system that can be solved with a trusted mediator can be solved by just the agents in the system, without the mediator, has attracted a great deal of attention in both computer science (particularly in the cryptography community) and game theory. In cryptography, the focus on the problem has been on secure multiparty computation, where each agent has some private information and the agents want to compute some function of this information without revealing it. This can be done trivially with a trusted mediator: the agents just send their private information to the mediator, who computes the function value and sends it to all of them. Work on multiparty computation conditions under which this can be done without a mediator, under the assumption that at most a certain fraction of the agents are faulty, and do not follow the recommended protocol. By way of contrast, game theory is interested in implementing mediators using what is called "cheap talk", under the assumption that agents are rational. We are interested in combining both strands: We consider games that have (k,t)-robust equilibria when played with a mediator, where an equilibrium is (k,t)-robust if it tolerates deviations by coalitions of rational players of size up to k and deviations by up to t players who can be viewed as faulty (although they can equally well be viewed as rational players with unanticipated utilities). We prove matching upper and lower bounds on the ability to implement such mediators using cheap talk (that is, just allowing communication among the players). The bounds depend on (a) the relationship between k, t and n, the total number of players in the system; (b) whether players know the exact utilities of other players; (c) whether there are broadcast channels or just point-to-point channels; (d) whether cryptography is available; and (e) whether the game has a (k+t)-punishment strategy; that is, a strategy that, if used by all but at most k+t players, guarantees that every player gets a worse outcome than they do with the equilibrium strategy.

This talk covers joint work Ittai Abraham, Danny Dolev, and Rica Gonen.No previous background in game theory is assumed.

Patricia Hu
Bureau of Transportation Statistics, U.S. Department of Transportation

"Transportation System Reliability: Challenges and Opportunities"

Everyday challenges will be presented to the participants of this year's Program on Games, Decisions, Risk and Reliability (GDRR) to explore solutions. The presentation will highlight the trends, potential contributing factors, and impacts due to the lack of reliability of the nation's transportation system. The discussion will touch on multiple transportation modes and from the perspectives of both moving people and goods. Other than capacity constraint and unanticipated events (e.g., crashes or inclement and extreme weather), the reliance of future transportation on automation and IoT would undoubtedly further impact system reliability in an unprecedented manner. An open dialogue on how GDRR can help ensure a reliable transportation system will conclude the presentation.

James Lambert
University of Virginia

"Disrupting Priorities of Engineering Systems"

This talk describes risk and resilience of engineering and enterprise systems to emergent and future conditions including natural and human induced hazards, technologies, regulations, behaviors, markets, demographics, supply chains, workforce, environments, etc. An emphasis is the quantification of risk, resilience, security, and trust as disruptions of systemic priorities. Examples will include a broadband wireless network for public safety, a maritime container port, airport runway safety, an energy grid of developing countries, and bidirectional chargers and microgrids for fleets of electric vehicles.

Bo Li
University of Illinois

"Gradient Boosting Trees for Spatial Data Prediction"

In recent years, statistical machine learning approaches have been extremely popular largely due to its superior performance in prediction. Of all the commonly used machine learning tools, the gradient boosting tree is usually the favored vehicle for many practitioners. On the popular data analytics competition platform Kaggle, gradient boosting is the winning algorithm for almost every structured data. Besides its superior prediction performance, the gradient boosting trees also enjoy the interpretablility of a non-parametric additive model and its fitting algorithm can be paralleled. In this project, we extend this powerful machine learning technique to the realm of spatial data analysis. The proposed approach does not require any parametric assumption on spatial correlations and enjoy all the advantages of gradient boosting. We illustrate the potential of the data with application on prediction of HIV new diagnose rates for all counties of the United States.

Feng Liang
University of Illinois

"Bayesian Regularization: Asymptotic Properties and Computation"

A central issue in statistics and machine learning is overfitting. In this talk, we introduce a general framework for effective regularization from a Bayesian perspective. In the proposed framework, Bayesian regularization is induced from scale mixtures of Laplace priors, including the regularization from spike-and-slab Lasso priors and the double Pareto priors considered in the Bayesian literature, as well as some known regularization considered in the penalization literature as special cases. The MAP (maximum a posteriori) estimator from our method gives rise to a new non-convex penalty approximating the L0 penalty. Our theoretical results show that the proposed Bayesian regularization enjoys optimal theoretical properties in terms of the L-infinity estimation accuracy for a large class of statistical models. For fast and efficient computation, EM algorithms can be employed to compute the MAP estimator. Our empirical studies confirm the theoretical findings regarding the attractiveness of the proposed Bayesian regularization. (The talk is based on joint work with Lingrui Gan from Facebook and Naveen N. Narisetty from University of Illinois at Urbana-Champaign.)

Igor Linkov
Carnegie Mellon University
Risk and Decision Science Team, US Army Engineer Research and Development Center

"Resilience: State of Science and State of Applications "

This presentation will review the history of risk assessment and management in the USA, discuss the emergence of resilience management, and the role of both constructs in addressing emerging risks. At the policy level, Resilience was a priority for Obama administration, especially in the context of climate change. Trump's administration is shifting the focus from climate change towards cyber and supply chain resilience, as it is reflected in recent Executive Orders. A major resilience impediment includes the lack of science of resilience, especially as it relates to assessing risks. Risk and Resilience are often used as synonymous even though they have a very different meaning, Risk-based approaches have been used to assess threats and mitigate consequences associated with their impact. Risk assessment requires quantifying the risk of failure for each component of a system and associated uncertainties, with the goal of identifying each component's contribution to the overall risk and ascertaining if one component poses substantially more risk than the others. These components become the basis of quantitative benchmarks for the system, and becomes the de facto standard for system improvements designed to buy down risk. In contrast to the definition of risk, resilience is focused on the ability to prepare and recover quickly from threats which may be known or unknown. Resilience is a property of the system itself and can be measured without identification and assessment of threats which act on or within a system. Managing for resilience requires ensuring a system's ability to plan and prepare for a threat, and then absorb, recover, and adapt. Coupled with a systems view that decomposes components across physical, information, cognitive, and social environments in which the system exists, is the basis of an approach to quantifying resilience with decision analytical tools and network science approaches.

I will present case studies of resilience assessment in the areas of infrastructure, transportation, cybersecurity, and organizational behavior using tools of decision analysis and network science. In all the cases, rapid technological evolution, combined with the unprecedented nature and extent of emerging threats defy us to enumerate all potential hazards, much less estimate reliable probabilities of occurrence and the magnitude of consequences. A comprehensive approach to protecting the nation's critical infrastructure, economy, and well-being must be risk based—not risk exclusive—and must provide a way for decision makers to make their organizational systems resilient to a range of threats within specific cost and time restraints.

George Michailidis
University of Florida

"Network Connectivity and Implications for Systemic Risk"

Interconnectedness amongst financial institutions has been implicated as a significant contributing factor to the 2008-09 crisis, where shocks were amplified to becoming systemic events. We study two types of networks: correlation network based on publicly traded bank returns, and a physical network based on interbank lending transactions and discuss various analytic approaches for studying their connectivity patterns over time. Some key findings include: (i) both networks behave similarly pin the period preceding the 2008-09 crisis, (ii) during the crisis the correlation network shows an increase in interconnectedness while the physical network highlights a marked decrease in interconnectedness. Moreover, these networks respond differently to monetary and macroeconomic shocks. Physical networks forecast liquidity problems, while correlation networks forecast financial crises.

Gordon Milbourn III
The MITRE Corporation

"Data Analytics and Gaming in the Prevention of Healthcare Fraud"

U.S. healthcare costs in 2018 were approximately $3.65 trillion, representing about 17.8 percent of gross domestic product. Healthcare claims fraud: is estimated to be between 3 percent ($110 billion) and 10 percent ($365 billion) of costs; drives up the cost of healthcare to legitimate recipients; is often perpetrated by organized criminal groups; and frequently entails the commission of other crimes such as identify theft, money laundering and tax evasion. Enormous amounts of data are regularly generated in the healthcare field, and powerful data analytics are important to identifying fraud indicators in that data. MITRE teams have identified new analytic approaches, such as billing in excess of 24 hours in a day when data from multiple payers is combined. Gaming has been shown to be very advantageous in identifying the risk that policy changes under consideration might create unanticipated consequences, which could include inadvertently opening the door to increased fraud. Because preventing fraudulent payments is far superior to paying and then chasing them, these approaches are vital to a successful organizational anti-fraud program.

Tao Pang
North Carolina State University

"Portfolio Optimization with Conditional Value-at-Risk"

We consider a portfolio optimization problem of the Black-Litterman type, in which the investor's view can be incorporated under a Bayesian framework. We use the conditional value-at-risk (CVaR) as the risk measure and the multi-variate elliptical distributions, instead of the multi-variate normal distribution, to model the financial asset returns. We propose an approximation algorithm and establish the convergence results. Based on the approximation algorithm, we derive a closed-form solution of the portfolio optimization problems of the Black-Litterman type with CVaR.

Nalini Ravishanker
University of Connecticut

"Modeling Approaches for High-Frequency Financial Time Series"

Analyzing high-frequency time series is increasingly useful with the current explosion in the availability of these data in several application areas, including but not limited to, climate, finance, health analytics, transportation, etc. This talk will give an overview of two statistical frameworks that could be useful for analyzing high-frequency *financial* time series leading to quantification of financial risk. These include a distribution free approach using penalized estimating functions for modeling inter-event durations and an approximate Bayesian approach for modeling counts of events in regular intervals. A few other potentially useful lines of research in this area will also be introduced.

Jesus Rios
IBM Research AI

"Adversarial Risk Analysis"

I will present in this talk the basic ideas leading to the current framework for Adversarial Risk Analysis (ARA) as it was conceived at SAMSI around 2006-07. In particular, I will discuss the differences between ARA and the Game Theoretic approach to model adversaries behavior. This presentation will be useful to set the stage for Thursday's working group session on ARA in which we will explore future research directions on ARA.

Bob Sivinski
Executive Office of the President

"The Foundations for Evidence Based Policymaking"

The Federal government has always sought to ensure that decisions are informed by sound evidence. Several recent developments across government will expand the capacity and capability for federal agencies to generate evidence that can improve effectiveness and efficiency of programs, improve the interface between the government and the public, and better understand risk, costs, and benefits.

Weidong Tian
University of North Carolina, Charlotte

"Dynamic Financial Decisions under Financial Risks"

In this talk, we discuss two kinds of dynamic financial decision - dynamic hedging of contingent claim and dynamic asset allocation - in the presence of concerns on financial risk and the implementation of financial risk management. For the dynamic hedging decision, we focus on the asset pricing implications of financial risks, such as liquidity risk and model risk. We also discuss some challenges of deep dynamic hedging of contingent claim. Then we discuss the dynamic asset allocation under several important financial risk management measures and formulate the equilibrium (game) under financial risks.

Bowei Xi
Purdue University

“A Game Theoretic Approach for Adversarial Machine Learning â€” When Big Data Meets Cybersecurity”

Nowadays more and more data are gathered for detecting and preventing cyber attacks. Unique to the cybersecurity applications, learning models face active adversaries that try to deceive learning models and avoid being detected. The existence of such malicious adversaries motivates the development of robust learning techniques. Game theory offers a suitable framework to model the conflict between the adversaries and the defender. We develop a game theoretic framework to model the sequential actions of the adversaries and the defender, allowing players to maximize their own utilities. For supervised learning tasks, our adversarial support vector machine has a conservative decision boundary, whereas our robust deep neural network plays a random strategy inspired by the mixed equilibrium strategy. On the other hand, in real practice, labeling the data instances often requires costly and time consuming human expertise and becomes a significant bottleneck. Our novel grid based adversarial clustering algorithm is able to identify the high confidence normal regions and the overlapping areas inside a mixed cluster, and identify outliers which may be potential anomalies.

Jun Zhuang
University at Buffalo

“Game Theory and National Security”

Hundreds of billions of U.S. dollars have been allocated to homeland security since 9/11/2001. How to optimally allocate these resources remains a challenging issue, especially considering the fact that adversaries are intelligent and adaptive. We will present a sequence of game theoretical models to identify equilibrium strategies for resource allocation to counter both terrorism and natural disasters. We will also discuss models of secrecy and deception, games between defenders, games between attackers, robust games, multiple-target games, and multiple-period games.