NERCCS 2019:
Second Northeast Regional Conference on Complex Systems

April 3–5, 2019 Binghamton, NY

Abstracts / Papers

Abstracts and Papers for the Second Northeast Regional Conference on Complex Systems

Complex Systems Society US Northeast Chapter
Center for Collective Dynamics of Complex Systems
Binghamton University, State University of New York
http://coco.binghamton.edu/nerccs/

April 3–5, 2019
Binghamton, NY

Contents

Part I
Keynote/Invited Talks

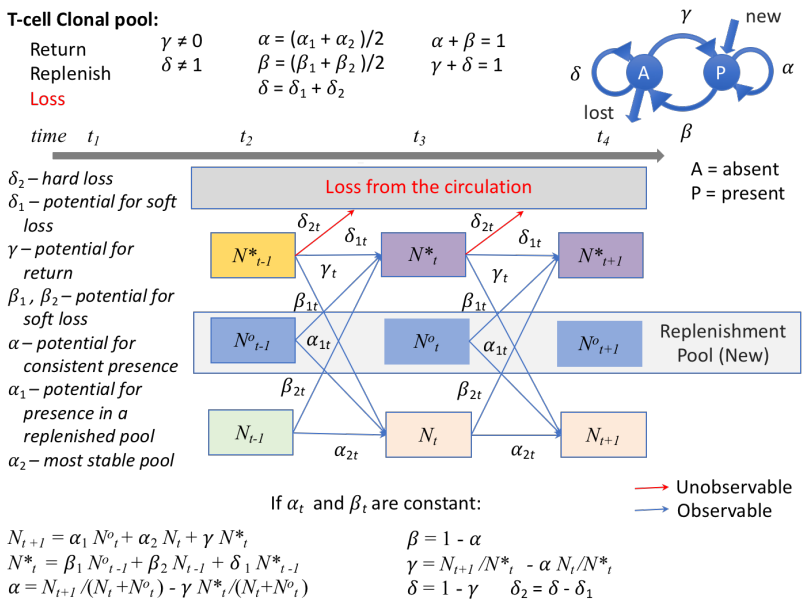
Understanding adaptive immune memory self-maintenance: a data-driven approach

Elena N. Naumova[1*], Yuri N. Naumov[2], Jack Gorski[3]

[1] Tufts University, Boston MA USA
[2] Smart Diagnostic Medica, LLC, Boston MA USA
[3] Blood Center of Wisconsin, Milwaukee WI USA
[*] corresponding and presenting author

Key words: data-driven approaches to complex systems; self-similarity and fractals; adaptive immune memory; adaptive systems and non-linear dynamics

Adaptive immune memory to pathogens can be described as a dynamic system that undergoes generation, maintenance, and immunosenescence. This system exhibits a scale-free behavior common in many living organizations at the micro- and macro-scales. The immune system operates on a large cellular component dedicated to the generation of adaptive memory to pathogens. It is this component of immunity which is the most instructive in understanding the life stages of humans. To observe and measure the formation, maintenance, and decline of adaptive immune memory to pathogens is costly and require longitudinal life-time assessments. We propose a data-driven approach to explore the stability of immune response to a common pathogen. The overall concept is based on a premise that in order to self-organize, an adaptive memory needs to be updated and a regular re-exposure to a common pathogen serves as a reminder. The recurrent exposures over life time inform and reinforce responses in a self-similar manner. Overall repertoire temporal stability, as defined by frequency of longitudinal occurrence, is governed by clonotype dominance and by a fraction of continuously circulating clonotypes. Our experimental data indicate that middle-aged adults exhibited the most stable and most complex repertoires. We model and quantify the maintenance of adaptive immunity by considering a Markov process that allows for a renewal pool (N^o), a sustained pool (N), and an influx pool (N^*), as shown in Figure below. We derived the initial parameters based on a set of longitudinal experiments. In the cohort of healthy donors, each encounter with an infectious agent was unique for every person. Yet, the commonality in responses formed "immuno-logical kinship" among all affected individuals, manifested by a relatively small, systematically preserved, dominant T-cell clonal pool, observed in every person. The findings from the proposed data-driven approach demonstrate how influenza-specific T cells participate in the dynamic process of system self-maintenance.



Part II

Wednesday April 3rd
4:30-5:50pm
Contributed Talks 1:
Biological Systems
(Symposium Hall)

Complex cell geometry when estimating glutamate transporter density in astrocytes

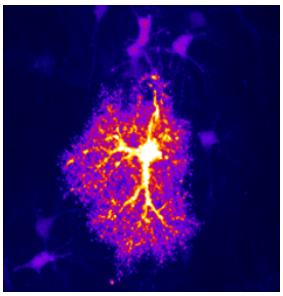
Cassandra Williams[1], Anca Rădulescu[1], Annalisa Scimemi[2]

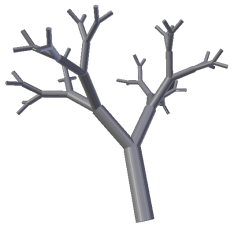
[1]Mathematics, State University of New York, New Paltz

[2]Biology, University at Albany

williamc18@hawkmail.newpaltz.edu

Glutamate is the main excitatory neurotransmitter released in the brain. Its removal from the extracellular space is important to terminate synaptic transmission between neurons, and prevent build-up of neurotoxicity. The removal process is intermediated by non-neuronal cells called astrocytes. These take in the excess extracellular glutamate via prism-shaped cross-membrane transporters densely expressed in the wall of the cell membrane. To understand their impact on neurotransmission efficiency, one needs to estimate the density of transporters for an average astrocyte [1, 2].





All existing computations are based on simplifying assumptions of spherical shape for a typical astrocyte. However, the actual, 3-dimensional fractal geometry of an astrocyte may drastically reduce this number, since cross-membrane transporters cannot collide. We use a geometric modeling argument, based on the known crystal structure of the transporter, to study how the structural complexity of astrocytic processes influences the surface density of transporters. One can then use Monte Carlo reaction-diffusion simulations to determine whether these theoretical estimates challenge our knowledge of how glutamate transporters shape efficiency of synaptic transmission.

References

[1] Jeffrey S Diamond. Deriving the glutamate clearance time course from transporter currents in CA1 hippocampal astrocytes: transmitter uptake gets faster during development. *The Journal of Neuroscience*, 25(11): 2906–2916, 2005.

[2] Annalisa Scimemi et al. Neuronal transporters regulate glutamate clearance, NMDA receptor activation, and synaptic plasticity in the hippocampus. *The Journal of Neuroscience*, 29(46): 14581–14595, 2009.

Complex Statistical Interaction in Biology

István A. Kovács[1,2,3], Albert-László Barabási[1,2,3,4]

[1]Northeastern University, Network Science Institute and Department of Physics, Boston, MA, [2]Dana-Farber Cancer Institute, Center for Cancer Systems Biology, Boston, MA, [3]Central European University, Department of Network and Data Science, Budapest, Hungary, [4]Brigham and Women's Hospital, Harvard Medical School, Division of Network Medicine and Department of Medicine, Boston, MA

Human genome sequencing lead to an extended landscape of genetic variation. Although each individual carries dozens of deleterious mutations, each of which should have a dramatic impact on our health, life carries on due to extensive genetic buffering. Therefore, interpreting genetic information requires the understanding of not only the impact of individual mutations but also their interactions. Genetic interactions occur when the combined impact of two mutations results in an unexpected phenotype, for example a positive interaction in the case of genetic buffering. Negative interactions are even more striking, such as synthetic lethality, where two individually mild mutations lead to cell death. Understanding and predicting (even higher order) genetic interactions is a key to better understand complex traits, missing heritability and genetic buffering in humans. Here we show that the network patterns of the already mapped interactions can be used to reliably predict so far undetected, yet biologically relevant functional interactions. First, we predict the existence and sign of genetic interaction in yeast, offering a remarkable, 95% sign precision at the genome-scale. Moreover, we are also able to address the combinatorial explosion and infer higher order genetic interactions between multiple genes, starting with genuine three- and four-gene interactions. Second, we apply our methodology to predict novel, clinically efficacious drug combinations as well as adverse drug-drug interactions. Our results can be applied more broadly, enabling us to better understand the emergence of biological function under both healthy and pathological conditions.

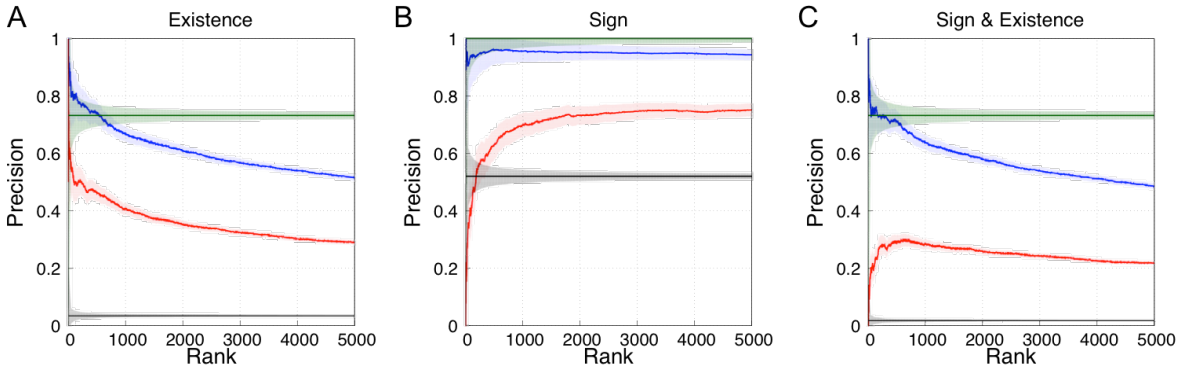


Fig. 1. Inferring genome-wide genetic interactions in yeast. The blue curve stands for our network-based approach, while the red curve illustrates a state-of-the-art method, based on the triangle-closure principle and social balance theory. We show the precision in a 50% computational cross-validation approach for the top 5000 predictions, using the data from Costanzo et al. *Science* (2016). **A**) Predicting the existence of a genetic interaction. **B**) Predicting the sign of a genetic interaction. **C**) Combined inferrence of the existence and sign.

Exploring Spillback and Dilution Effects: A Unifying Framework Based on an Emergent Pathogen in Bumble Bees

Brendan Case[1] Laurent Hébert-Dufresne[1]

[1]Department of Computer Science and Vermont Complex Systems Center, University of Vermont, Burlington VT, USA

With the rise of Emergent Infectious Diseases in a variety of human and animal populations, there is a need to understand the effects of disease spillover from reservoir populations onto new susceptible hosts. Using as a case study the spillover of Deformed Wing Virus from honey bees to bumble bees through shared flower pollination, we present a simple framework in which the presence of separate susceptible populations leads to an increased risk in disease outbreak (see Fig. 1). Our mean-field approach is simple enough to study analytically, yet able to fit observations from an artificial inoculation experiment.

In addition, we add a spatial component to the well-mixed model by building a tripartite network using real-world floral landscapes. Bees are connected to flowers based on the proximity of their colony to the flower patch, while taking into account other terrain features.

Both our models allow us to explore two important concepts from theoretical epidemiology: spillback, in which the introduction of a susceptible population leads to increased infection levels, and dilution, where the addition of a population serves as an infection sink, thereby decreasing overall outbreak intensity. While there is evidence supporting both these hypotheses in various settings, these two effects are sometimes framed as opposing outcomes. However, in showing that flower prevalence and diversity causes dilution, while presence of separate bee populations causes spillback, our results support the idea that this is a false dichotomy in realistic community assemblages.

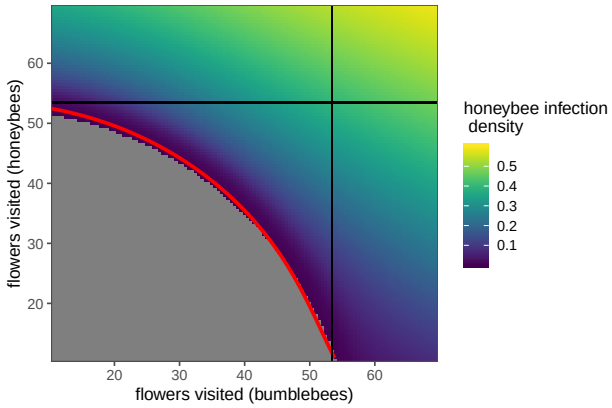


Fig. 1: Infection density in honeybees after simulating the mean-field model until convergence. Each axis corresponds to the number of flowers a bee visits per day. The red line shows the reproductive threshold (i.e. where $R_0 = 1$) predicted by our derived R_0, while black lines show the thresholds for honey/bumble bees if the presence of the other population were ignored. In a large part of the parameter space the infection was still present, even when the outbreak would die out in both populations if they were separated from each other.

Flowers as dirty doorknobs: Virus transmission through flowers depends on floral diversity

P. Alexander Burnham[1,2], Samantha A. Alger[1,4], Humberto Boncristiani[3], Laurent Hebert-Dufresne[2]

[1] Department of Biology, University of Vermont, Burlington, VT
[2] Complex Systems Department, University of Vermont, Burlington, VT
[3]Entomology and Nematology Department, University of Florida, Gainesville, FL
[4] Environmental Scientist, Vanasse Hangen Brustlin, Inc., S. Burlington, VT

Evidence is mounting that RNA viruses, likely originating from honeybees, have been spilling over into wild bee communities, potentially through shared flowers. In the field of disease ecology, the dilution hypothesis examines the idea that diversity can dampen the effectiveness of a transmission route. In this study, we experimentally demonstrate how flowers may be facilitating the spillover of Deformed Wing Virus (DWV) from honey bees to bumble bees. In addition, through a combination of laboratory experiments and mathematical modeling, we examine potential mechanisms for dilution through increased floral diversity. Clean bumble bees were allowed to forage on honeybee-infected flowers. We found that DWV can be picked up by bumble bees and that increasing foraging time increased viral load. To simulate this transmission route, we used experimentally estimated parameters in a cellular-automata model. Infected bees deposited DWV on a flower with a probability, γ and infected flowers visited by susceptible bees picked up the virus with the probability, β. Distributions of flower species constructed using data from field surveys were examined with their corresponding γ estimates. Additionally, the vital statistics for honey bees, bumble bees and flowers were estimated and incorporated. Our model showed that the parameter estimates for viral deposition and probability of infection are able to drive a simulated spillover event mirroring empirical data from field surveys. The inclusion of floral diversity (modeled as a distribution of γ) dampened overall disease prevalence. Future work should look to optimizing floral species compositions to maximize floral diversity and minimize disease spillover.

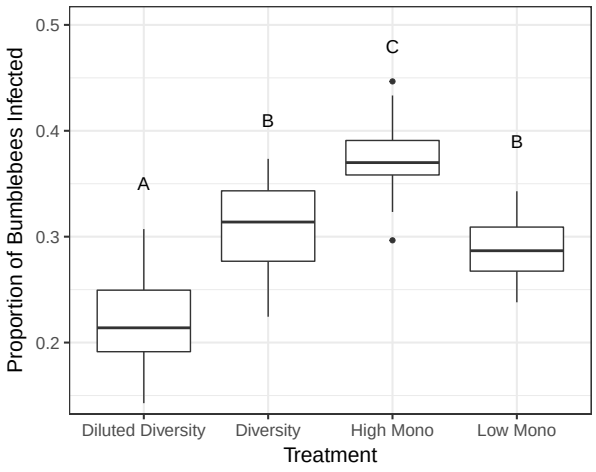


Figure 1: The proportion of bumble bees infected during simulation for a mixture of diverse virus harboring flowers and non-harboring flowers (Diluted Diversity), Diverse harboring flowers (Diversity), a monoculture with a high probability of viral harboring ($\gamma = 0.35$) (High Mono) and a low monoculture ($\gamma = 0.22$) (Low Mono). Letters represent the results of a Tukey post-hoc test.

Part III

Wednesday April 3rd
4:30-5:50pm
Contributed Talks 2:
Dynamical Systems
(Engineering Science 2008)

Least Action Principle Applied to a Non-Linear Damped Pendulum

Katherine Rhodes and Ashwin Viadya[1]

[1] Department of Mathematical Sciences Montclair University, State University of New Jersey
vaidyaa@mail.montclair.edu

Abstract

The principle of least action is a variational principle that states an object will always take the path of least action as compared to any other conceivable path. This principle can be used to derive the equations of motion of many systems, and therefore provides a unifying equation that has been applied in many fields of physics and mathematics. Hamilton's formulation of the principle of least action typically only accounts for conservative forces, but can be reformulated to include non-conservative forces such as friction. However, based of the work of Wang and Wang,[1] for sufficiently large damping parameters and the right model, the object will no longer take the path of least action. Through numerical simulation, this is shown to be true for two systems, an object in free fall in a fluid and a non-linear damped pendulum. For the pendulum, both linear and cubic damping are explored.

Figure 1 shows the optimal action as a straight line for two different values of damping for a cubically damped non-linear pendulum. This is compared with the action of 100 numerically generated nearby paths, denoted by the blue stars. In this example, we are taking very small perturbations of the true path. Figure (a) has a damping coefficient of 100, Figure (b) has a damping coefficient of 500. For small values of damping the principle of least action holds. However, as damping becomes large, this is no longer the case.

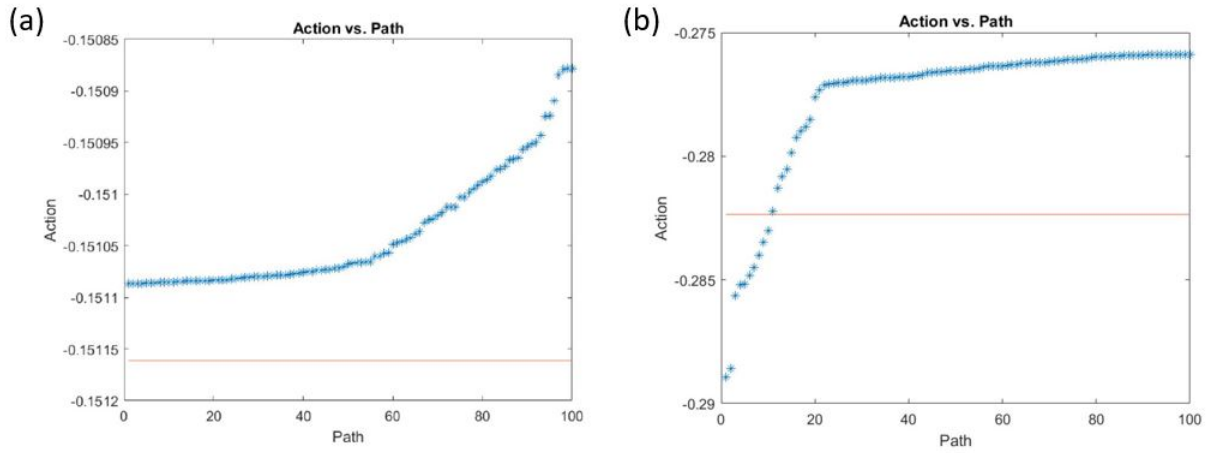


Figure 1: Action of the true path of a cubically damped pendulum (red line) compared to the action of 100 randomly generated variation of the true path (blue stars) for damping coefficients of (a) 100 and (b) 500. The x-axis represents the index for each path and the y-axis denotes the action of that path.

[1]Wang, Qiuping A., and Wang, Ru. "Is is possible to formulate least action principle for dissipative systems?" (2012).

On the cusp: Bistability and attractor strength predict reaction time Hurst exponents

P. Adrian Frazier
Center for the Ecological Study of Perception and Action
University of Connecticut

Research examining fluctuations in reaction time (RT) data on simple cognitive tasks consistently returns evidence that such fluctuations are long-range persistent and self-similar, as measured by the Hurst exponent (*H*), where higher values index greater self-similarity. *Voluntary* performance, which relies less on experimental stimuli for response, has higher *H* than that which is *compelled*, which depends more on experimental stimuli. For instance, Kello et al. (2007) presented participants with sequences of "<" and ">" one at a time over the course of more than 1000 trials. In one condition, the sequences were predictably ordered (<, <, >, >, <, <, >, >,…) and in another condition, they were randomly ordered. RT timeseries in the former condition returned higher *H* than RT series in the latter. Although results like this have been quite reliable, distributions of *H* tend to be highly variable, suggesting that the voluntary-compelled story is incomplete. Catastrophe theory and synergetics provide clues about where to go.

Dynamic systems with stable patterns of behavior, even with fluctuations, are governed by *attractors*. The stronger the attractor, the more rapidly it dampens fluctuations, making the pattern less variable. *Control parameters* govern the properties of a system's attractors. *Critical* values of a control parameter occur where a pattern becomes destabilized and a new attractor takes over, switching the system from one pattern to another. Systems near criticality exhibit long-range persistence: a perturbation at one point has consequences that are felt over a long range due to the weakened attractor. Through a series of computer simulations modeling the *cusp catastrophe*, I have verified that self-similarity is also systematically related to criticality and the strength of an attractor. The system has two control parameters, one of which, the *bifurcation parameter* β, determines the strength and number of attractors. The system has one attractor for low β, another for high β, but two attractors for a region in between (the *hysteresis* region). Among other things, the results show a nonlinear increase of *H* toward 1.0 as β approaches a critical value.

Empirically, these results suggest that timeseries resulting from an unstable mode of behavior should have lower *H* than that which is stable. To test this, I examined data from a follow-up to the Kello et al (2007) study. The sequence was 8 arrows long, and participants either memorized the sequence ahead of time or did not. Two strategies were available for those who either memorized or discovered the sequence: 1) react to the cue or 2) anticipate the cue. Both strategies are evident in a bimodal distribution of RTs. Figure 1 shows four RT series with a sliding window *H* (black, with window trailing in orange/purple) and accreting *H* (increasing number of time points) in red. Whether strategy 1 or 2, when stable from beginning to end, *H* remained relatively low (top row). Strategy 2 may have a weaker attractor than 1. Switching back and forth and switching from one to the other results in much higher *H*. Interestingly, in the regions to the left and right of the transition, performance is relatively stable, and likewise, *H* is much higher in the transition zone than in the stable zones.

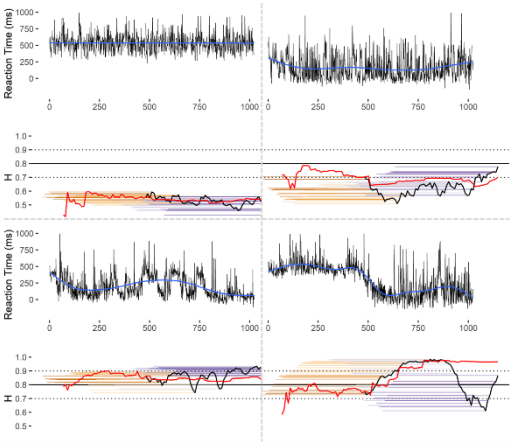


Figure 1: Time series, moving window H (black with window trailing in orange/purple), and accreting H (red).

Entropy Production and Segré-Silberberg effect

Joseph Pateras, Bong Jae Chung, and Ashwin Vaidya
Department of Mathematical Sciences, Montclair State University, Montclair, NJ 07043
paterasj1@montclair.edu

The Segré-Silberberg effect, first observed by its eponyms in 1961[1], has hence become the subject of a large body of rigorous fundamental and applied research. The first experimental observations of this natural oddity showed that sedimenting spheres suspended in Poiseuille flow tend to settle at a position within a channel equal to about $0.6 - 0.7$ times the radius of the channel. Ever since, researchers have aimed to explain the phenomenon that produces the Segré-Silberberg radial equilibrium displacements in all types of flow parameter regimes. Many such research projects have described a 'wall effect' felt by particles near the boundaries of flow. Analytical[2] approaches have been used to show the potential for varying equilibrium positions. Numerical[3] simulations have described a 'wall-effect' and quantified lift force, drag force, settling distance, and other interesting quantifiers in the extensively studied near-wall interactions. Further experimental studies have found reality always agrees with the aforementioned results, even with many sedimenting particles of varying particle-to-channel size ratios[4]. All the above studies have reached the consensus that the settling distance of sedimenting particles in many flow regimes contain an equilibrium position somewhere near the phenomenal Segré-Silberberg radius.

In the search for some underlying catalyst for these observations we attempt to employ novel tools and quantify properties of this particular dissipative system that have yet to be considered. The purpose of this study is to take a look at the Segré-Silberberg phenomenon from a thermodynamic perspective. Our study probes whether or not the maximum entropy production principle (MaxEP[5] principle) is able to predict the peculiar organization of the Segré-Silberberg effect. The MaxEP principle has been invoked in many realms of complex systems and self-organization as a stability catalyst[6].The principle states that a system, out of thermal equilibrium, will tend to a state in which it produces entropy at the highest rate. In order to apply this method to flow past a circle, we employed COMSOL to model a single suspended particle in flow near a wall. However, the sedimenting object in our model is stationary as water flows by it, simulating buoyancy. Using this model we can calculate some thermodynamic properties of the system; chief among them, the rate of entropy production (REP). Calculations for REP are carried out using the deformation function. Preliminary results have suggested for a particle within a flow of relatively low Reynolds Number placed anywhere within a channel REP attains a steady value as the flow matures. For higher Reynolds Number ($Re > 60$) REP oscillates around a steady value in time. Furthermore, it is shown in our results that the radial positioning of the sedimenting particle causes the system to reach a peak REP when the particle is placed near the Segré-Silberberg radius.

These results bode well in an attempt for a thermodynamic explanation for the formation of the Segré-Silberberg equilibrium position. Our results suggest that the pattern formation observed regularly in experiment and simulation could be attributed to the system's search for a state corresponding to the maximum rate of entropy production: agreeing with the MaxEP extrema principle.

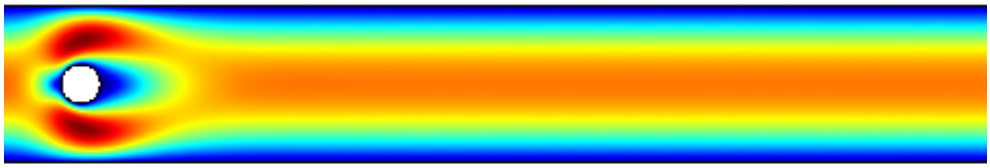


Figure 1: COMSOL model depicting velocity field in 2-D flow past a circle. Re=5

References

1. Segre G; Silberberg A (1961). "Radial particle displacements in Poiseuille flow of suspensions". Nature. 189: 209210.

2. Mitchell H. W., Spagnolie S. E. (2015). Sedimentation of spheroidal bodies near walls in viscous fluids: glancing, reversing, tumbling, and sliding. J. Fluid Mech. Vol.772 pp. 600-629 Cambridge University Press.

3. Singh, P., Joseph D.D (2000). Sedimentation of a Sphere Near a Vertical Wall in an Oldroyd-B Fluid. Journal of Non-Newtonian Fluid Mechanics. pp. 179-203.

4. Morita Y., Itano T., Sugihara-Seki M. (2017).Equilibrium radial positions of neutrally buoyant spherical particles over the circular cross section in Poiseuille flow J. Fluid Mech. Vol.813 pp. 750-767 Cambridge University Press.

5. Chung, B.J., Vaidya, A. (2014). On the affordances of the MaxEP principle. The European Physical Journal B. 87: 20

6. Chung, B.J., Ortega B., Vaidya, A. (2017). Entropy production in a fluid-solid system far from thermodynamic equilibrium. The European Physical Journal E. 40: 105

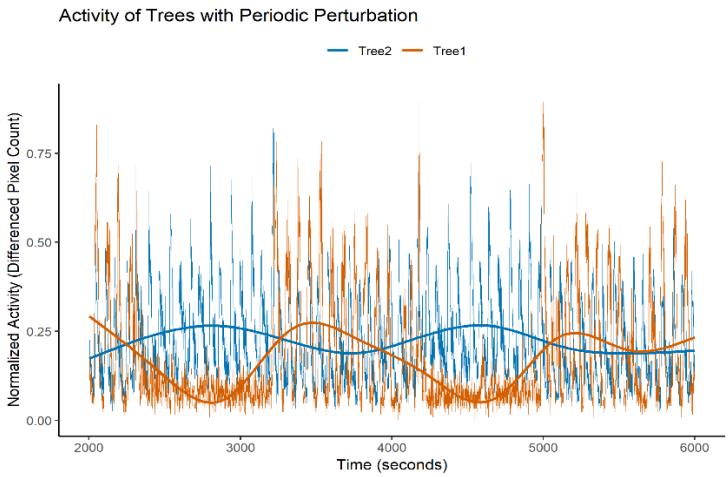
Coupled Dissipative Structures: A Physical Analog of a Coordinative Structure

Ben De Bari, James A. Dixon
Center for the Ecological Study of Perception and Action
University of Connecticut

In the early 20th century, Nicolai Bernstein observed the remarkably fluid production of a skilled task, despite the inordinately many degrees of freedom that must be rallied in its service. Somehow, Bernstein posited, these vast degrees of freedom come to be mutually constrained, and from such complexity emerges the simplicity of the intended act. Such a phenomenon is known now as a 'synergy' or 'coordinative structure'; the task-oriented, self-organized, mutual constraint of physiological degrees of freedom.

Kelso et al. (1984) provided a compelling experimental paradigm for revealing a coordinative structure. While participants repeatedly produced a phrase, an intermittent perturbation was delivered to part of the jaw. He observed adaptations in the speech apparatus remote from the affected area (e.g. the upper lip in response to a force on the lower jaw), such that the appropriate speech sound is produced. These results show that in a speech act, physiological degrees of freedom are mutually constrained in a task-specific way.

To provide insight into coordinate structures in living systems, we study dissipative structures, a general class of non-equilibrium self-organizing systems, of which living systems are a subset. Dissipative structures provide a simple system of the same physical class in which to explore the phenomenon. We study an electrical dissipative structure, a set of metal beads in shallow oil which, when subject to high voltage, spontaneously form into tree-like structures which maintain their structure and move about in the dish. Previous work showed that this system self-selects for states of increased rate of entropy production (morphologically and behaviorally), and thus can be cast as end-directed. These trees are coupled through their mutual influence on the distribution of charges in the dish and share a 'task-space' of maximizing the rate of entropy production.



We performed perturbation experiments akin to Kelso et al.'s speech study. A magnetic force was introduced to one tree structure, freezing out its mobility and reducing its ability to draw current. Consequently, other tree in the dish ramped up its activity, and the current through the whole dish remained stable.

This system offers a unique methodology to investigate coordinative structures, the form of their coupling, and the nature of the constraint imposed by the task. Given it's simple (relative to biology) processes, its behavior is readily open to empirical manipulation and physical explanation. The properties of this minimal model will ideally be generalizable to the actions and architectures of biology.

Part IV

Thursday April 4th
11:10am-12:30pm
Contributed Talks 3:
Data Science Special Session I
(Symposium Hall)

Deep Learning Based Cancer Classification of Pathology Slides using Cancer Cellularity Score of Pathology Patches

Kayvan Tirdad[a], Alex Dela Cruz[a], Hossein Rahnama[b], Alireza Sadeghian[a]
Tirdad@ryerson.ca, adelacru@ryerson.ca, hossein@ryerson.ca, asadeghi@ryerson.ca

[a]Dept. of Computer Science, Ryerson University, [b]RTA, Ryerson University

The analysis of pathology slides in the diagnoses of cancer tumors is commonly performed manually by expert pathologists. Starting at a low magnitude, a pathologist will analyze the entire pathology slide to identify patches of interest. Patches of interest are then analyzed at higher magnitudes, commonly 20x, as indicators to classify the slide. The current process is time consuming and does not necessarily yield consistent results across different pathologists. We propose a deep-learning based system that automatically classifies pathology slides using the cancer cellularity scores from the patches. Our system consists of two learned sub-systems, where the first system calculates cellularity score of each patch of the slide, and the second detects the patches of interest from low magnitude slide images. A combination of these sub-systems detects patches of interest and scores them properly to make a final decision regarding cancer detection in various tissues. Many patches are generated per slide during segmentation whereby an effective method to calculate the cancer cellularity score is trained. We modified inception-v4 neural network, to a deep regression model, to predict the cancer cellularity score of a patch.

The dataset consists of 2,579 sample patches extracted from 69 whole slide images that have all been labeled by an expert pathologist. The pathologist has assigned a tumor cellularity score to each patch in the range from 0 to 1. The dataset was split into training and test datasets containing 2,394 and 185 patch samples respectively. To further increase the number of sample size such that the neural network model may better learn some of the underrepresented labels, we have applied rotation and mirroring transformation on each patch, and hence generating total of 19,152 possible samples. The result of our model with five different configurations in terms of weighted normalization, mirroring and batch size presented in Figure 1. Since manually labelling all patches

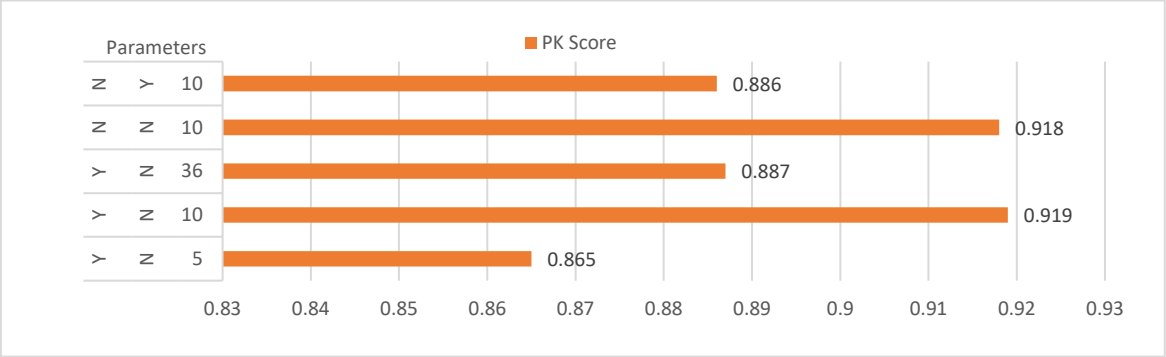


Figure 1: Evaluation results of inception-v4 model using various parameters. The prediction probability score is evaluated for each model on various parameters (weighted normalization, mirror and batch size).

of a single slide is time and resource consuming, our trained inception-v4 model can be used to label all the patches within a slide. However, this method can pose a computation challenge to label all the patches in a timely manner. As such, a dataset of whole slides with unlabeled patches are labeled using the model to produce a new dataset of labeled patches. The whole slide dataset and the new labeled patches is then utilized to train a generative adversarial network to more effectively generate the score. The combined cancer cellularity score of each patch in the slide is then feed into a classification model to classy the cancer grade of the slide.

Biomedical image analysis and cell localization via persistent homology

Ulgen Kilic[1], Michael Vaiana[2], Ethan M. Goldberg[3,4], Sarah F. Muldoon[1,2,5]

[1] Department of Mathematics, University at Buffalo, SUNY, Buffalo, NY
[2] CDSE Program, University at Buffalo, SUNY, Buffalo, NY
[3] Departments of Neuroscience and Neurology, The Perelman School of Medicine at The University of Pennsylvania, Philadelphia, PA
[4] Division of Neurology, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA
[5] Neuroscience Program, University at Buffalo, SUNY, Buffalo, NY

Biomedical imaging is a powerful tool for studying biological systems at the cellular level. In these images, it is often necessary to identify regions of interest (cells or nuclei) before further analysis can take place. Here, we present a new algorithm that applies persistent homology to biomedical images and uses persistence information to filter the results to identify locations in the image that are likely to correspond to cells. We apply the method to images in which cells appear either as filled regions (using 0-dimensional homology) or as donut-shaped objects (using 1-dimensional homology). Importantly, this method is robust to small perturbations, and therefore advantageous for use in images with small amounts of noise. We verify the performance of our methodology to identify cellular nuclei using the 2018 Kaggle Data Bowl Competition data set and provide example applications to neuronal calcium imaging data and myelin segmentation in microscopy images.

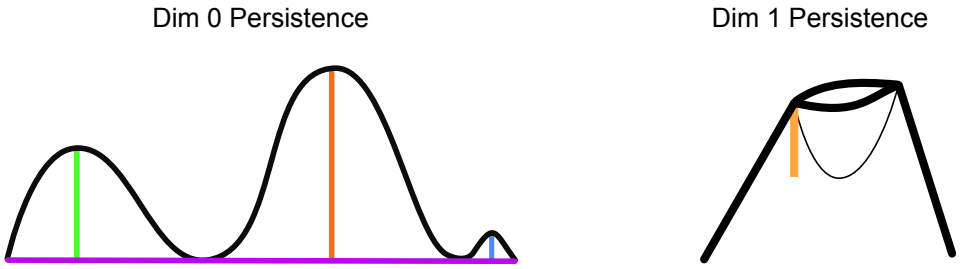


Figure 1: Persistence in images. A visualization of the information captured using persistent homology in images, where the y-axis represents the pixel intensity of the image. Thinking of the image as a function from the plane to $\mathbb{R}$, we can think of bright spots in the image as hills or peaks and bright closed loops as the rim of a volcano. *Left.* The 0-dimension persistence measures how large a peak is relative to its surroundings. When peaks are high and separated from other peaks by relatively deep valleys they will have large persistence. The lines drawn down from the peaks indicate the length of persistence. *Right.* The 1-dimension persistence measures how long closed loops last relative to their surroundings. A bright loop can be thought of as the rim of a volcano and its persistence length is how far the rim is from the basis of the lava.

Swarm Optimization in the Search of Extraterrestrials by Radio Signal Detection: Adaptive Strategies on Radio Telescope Data

Daniel K. Trembley
System Science & Industrial Science Engineering
Thomas J. Watson School of Engineering & Applied Science
State University of New York at Binghamton, NY - 13902, USA

Abstract

The idea that extraterrestrial civilizations (EC) exist has been theorized and speculated about since the time of the Copernican Revolution. The subject in the centuries that followed has mainly been left to the imagination of philosophers and science fiction writers. With the advent of radio astronomy and deep space satellite imagery in the 20th century, detecting an EC in the vast cosmos becomes a real possibility. Using the data from radio telescope arrays and space telescopes, the research focuses on regions within the Milky Way galaxy where exoplanets truly exist and where EC is most likely to reside. Since the Kepler Space Telescope (HST) has found Earth like analogues orbiting distant host stars within habitable zones of the exoplanet host star, and the SETI Institute has amassed radio signal data from those regions, the direction to where an EC might exist has never been more probable. The issue is how to analyze the massive radio data and classify the thermal and synchrotronic emissions in the radio transmissions into categories of expected signals and noises. This research explores the idea of using meta-heuristic techniques of Swarm Intelligence to filter, detect, and locate radio emissions sources.

Keywords: SETI, Radio Astronomy, Massive Data, Swarm, Optimization, Cockroaches

1. Introduction

The mostly likely source of the question, “are we alone in the universe?” must have been asked around the time of the Copernican revolution. When humanity lost their spot at the center of the universe, the paradigm shifted, and a new phase of scientific questioning began. Today’s optical and radio telescopes reach back across the universe, not just in distance, but also in time. The largest search space in all the universe must be the universe itself [1, 2], and all the data possible exists within this space. The Hubble Space Telescope (HST) observes deep space objects reaching distances of 13 billion years. The universe is essentially a time machine of our deep and near past [3], and allows the prognostication of the future by physical properties alone.

The question today is still, “are we alone in the universe?” The question must be understood within the scientific frameworks and methods of knowledge and understanding: why are we asking and what determines the expected answer. Anthropomorphizing the universe could be as detrimental to our own state-of-the-art science as religion was to science in the 15th and 16th centuries, in the time of such notables as Bruno, Kepler, and Galileo. Each faced their own dilemma of established beliefs as they challenged the dogmas of their times [4]. Therefore, the question really becomes one of constructing a cosmology that references our assumption and binds with testable theory. The logic is profound, and that everywhere in the universe appears the same should be telling. There are more of us out there.

2. Literature Review

The Drake equation [5, 6] is a probabilistic argument used to estimate the number of active, communicative extraterrestrial civilizations in the Milky Way galaxy. The number of such civilizations N, is assumed to be equal to the mathematical product of

1. Average rate of star formations (R_*) in the Milky Way galaxy
2. Fraction of formed stars that have planets (f_p)
3. Average number of planets that can potentially support life (n_e)
4. Fraction of those planets that actually develop life (f_l)
5. Fraction of planets bearing life on which intelligent, civilized life, has developed (f_i)
6. Fraction of these civilizations that have developed communications technologies that release detectable signals into space, (f_c)
7. Length of time over which such civilizations release detectable signals (L)

The equation is written as follows:

$$N = R_* * f_p * n_e * f_l * f_i * f_c * L$$

Using the probabilistic argument then leads to the next question, "how do you find a civilization in the vastness of the cosmos?" There are several issues with this type of research [7, 8]. The universe is a time machine. Seeing a star that is 50 light years away is looking at the star as it was 50 years ago, not how it is today. The true position of the star is only a projection of how gravity and space-time move, and the lifetime of stars is anywhere from a few million years to many billions of years. This makes the constituent parts of the Universe a spectrum from the very young to the very old.

The scientific research of radio astronomy traces its roots to pre-World War II technology. With the inventions of RADAR (Radio Detection and Ranging) for detecting enemy ships and aircraft, the engineering and science of RADAR sensor systems did not take long to repurpose to navigation, weather, and astronomy. Karl Jansky, the father and founder of radio astronomy, made three discoveries: that radio noise was from nearby thunderstorms, distant thunderstorms, and static from outer space [12]. To his credit, his name is the unit of measurement for the apparent brightness of a radio source [4, 12].

For about 60 years, radio astronomers have been listening to the cosmos, and not for the purposes of finding Little Green Men (LGM) [12, 19]. Many discoveries in modern astronomy and astrophysics use radio telescope data. Pulsars, Cosmic Microwave Background Radiation (CMBR), nebula chemical composition, etc., are just some of the phenomena that this data reveals [12]. Because the search for an EC transmission can come from any direction, this furthers the problem of detection. Pulsars are numerous and other measurements, such as optical telescope, have observable location properties. Since a radio telescope array is required to point in the direction of transmission at the exact time and frequency of the passing signal, the search for EC becomes even less probable of detection. Scientists only have a limited coverage range against the infinite sky [9, 10]. Additionally, noise from celestial emission sources in the frequency bands that an EC would be using also greatly affects probability of detection. Too much search space and too little signal -- the search for a voice in space goes unanswered, and the term the "Great Silence" seemingly confirms the Fermi Paradox [1].

The original estimates in the 1960s using the Drake Equation postulated that as many as 10,000 to 100,000 extraterrestrial civilizations exist within the Milky Way Galaxy. In 2009, the KST, detecting exoplanets through a technique called occultation [3], grew the estimates from 100,000 to 1,000,000. Since then, the Kepler Space Telescope (KST) has almost confirmed that every star has planets. The current confirmed

count of exoplanets from KST is 4,496 candidates, 2,337 confirmed, and 30 small and in the habitable zone [5, 8]. In a sense, target exoplanetary systems have been located, which methodically refines the search space parameters, so the search becomes purposely directed and not random or accidental. However, the amount of time listening and the massive data generated, presents the next dilemma for the researcher. That is, how to detect a signal when the signature of that signal is not known. The simple answer is to detect all signals, and categorize the known signals as false alarms. Anything not categorized is a candidate needing further investigation.

This research uses a modified version of the Particle Swarm Optimization (PCO) called the Improved Cockroach Swarm Optimization (ICSO). The ICSO uses the advantage of a "hungering" component to improve the search and population diversity [15]. The ICSO is membership to a family of heuristic algorithms known as Swarm Intelligence (SI) [14]. The SI methods work by decentralizing multi-agent environments, in this case, creating a virtual cockroach colony that clusters and disperses throughout the search space [14, 15]. The analytics of SI increases solution performance by identifying traits within a population and efficiently identifying local optima within the solution space.

3. Methodology

The SETI data set chosen for this research is from observations of Kepler 144b. Although unlikely that any advanced civilization lives there, Kepler 144b is within the habitable zone of its host star, though the planetary mass is not well understood. The time to orbit the host star is about 6 days, as compared to the Earth taking 365 days around the sun. The KST has also found a second planet within the system known as Kepler 144a, which could increase the chances of an EC. Because of the great distance to the Kepler 144 system, even if an extraterrestrial civilization existed, it would take over 1200 years to communicate one way [5, 8]. Any technologically advanced civilization in the Kepler 144 system that could receive radio transmission would not receive any signals from Earth until after December 24, 3106.

The telescope observations and data collection is from the Allen Radio Telescope Array located in Mountain View, California [5, 11]. Figure 1 shows a typical radio telescope. An array of these telescopes could consist of around 20 or more depending on the frequencies being researched. The more telescopes in the array, the better the resolution to the transmitting source [12]. Figure 2 depicts the radio telescope collection region and is an actual "radio" picture of the region where Kepler 114b resides [6]. The bright dots are stars, and the mottled dark background is the cosmic background radio sources.

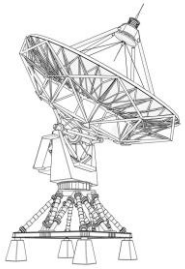


Figure 1 - Radio Telescope

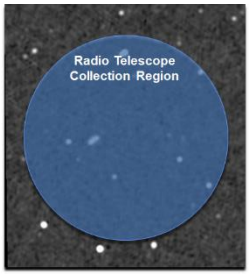


Figure 2 - Radio Telescope Collection Region

In Figure 3, the waterfall graph of the Kepler 114b at a frequency of approximately 3.127 MHz demonstrates frequencies being captured with respect to time. The graph has a semi-straight lighter color line against the dark backdrop. This indicates a transmitted source against the background noise of the sky, a local minimum within the search space of the data. However, the signal, as shown, is a synchrotronic emission source from a nearby nebula and not from Kepler 144b. Because of the regularity and the slight slope going left to right and from top to bottom, the signal source is not an EC [3, 12]. Since radio signals

have the multidimensional components of frequency, time, and power, and because radio telescopes range in many different sizes, the data capture of large meter wave length emissions from the electromagnetic spectrum needs to be determined in advance [9, 16].

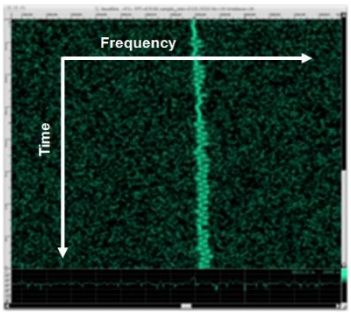


Figure 3 - Waterfall Graph

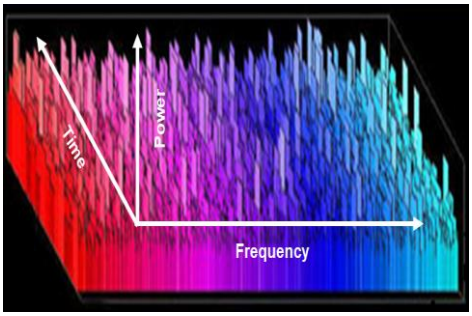


Figure 4 - Time Frequency Power Graph

Figure 4 shows a broadband emission captured. The time dimension is from current to past moving front to back, and the frequency dimension is from low to high moving left to right. The power dimension is from low to high moving in the vertical direction. Power is the energy of the radio emission, i.e., intensity, measured in a billionth of a Watt. An EC would likely be transmitting at a billionth of a billionth of a Watt [12], so the complexity of finding a signal in the data is difficult because of the sensitivity to lots of noise from multiple emitting sources, i.e., nebula, Earth's ionosphere, etc.

The SI method used is the Improved Cockroach Swarm Optimization (ICSO). Essentially, the adaptive technique of cockroach behaviors is applied [14, 15] to search the radio data. In Figure 5, the borders of the figure define the boundaries of the search space, as the ICSO would observe and be confined to it. The dots on the left-hand side of the search space are the cockroaches. The cockroaches are in their swarm dispersion configuration that is both uniform and randomly. These cockroaches are in search of a signal detected from a radio source in the time-frequency domain of the recorded signal. The green circle at the right-hand side of the search space represents the emission source of interest within the time-frequency band.

Figure 5 - ICSO Frame One

Figure 6 - ICSO Frame Two

Figure 7 - ICSO Frame Three

Figure 8 - ICSO Frame Four

In Figure 6, the cockroaches are moving about the search space as a randomly clustered group. The dispersion pattern changes; however, the swarm is still collectively intact and the group moves toward the target, "Signal from Star." Because some of the outlier cockroaches are detecting a gradient within the search space, a non-normalized pattern of difference, the swarm collectively moves towards but not beyond the outliers.

The collective dispersion of the cockroaches gets tighter and descends on the perceived local optima, Figure 7. This is the "food" source, or the candidate signal from the EC. The idea is to find the frequency source, say a transmission at 21.175Hz, that is just outside the emission band of molecular hydrogen (H1). In this case, the entire swarm collapses on the transmission source. This identifies a region of interest (ROI), and in the final phase, the cockroach collective is entirely on the source, Figure 8.

The search process is repeated through each time-frequency domain by redistributing the cockroach colony randomly within the search space of the nearest neighboring frequency. A greater pattern, shown in Figure 3, emerges. The caution here is to properly classifying the radio source. As radio signals travel through space, their frequencies change based on objects encountered, such as dust and gas. This is a physical property also called dispersion. The cockroach swarm, through the time-frequency domain of gradually changing frequencies of the same source [17], can classify the known patterns. It is when the changing frequencies become perturbed and draw attention.

4. Experiment

The data from the SETIQuest website consists of six compressed files of 2-GB each with additional smaller size descriptor header and META files. The uncompressed data is approximately one terabyte. The ICSO algorithm is written in the C programming language using Microsoft Visual Studio 2015. The pseudocode algorithm provided by the original ICSO article [14] is used. The ICSO algorithm defines the following variables in Table 1.

Table 1 - ICSO Variable Definitions

Variable	Definition
N	Number of Roaches
x_i	Cockroach Position
step	Constant Define
p_i	Personal Best Position
p_g	Global Best Position
D	Dimensional Random Vector
x_k	Random Position
w	Inertial weight
x_{food}	Position of Food
T	Hunger
ct	Migration speed

A scaled down version of the first 100-megabytes, or approximately one-hour of observation for the initial phased experiment is used, and a sensitivity of 1/10 of a billionth-watt will indicate a potential gradient. The first step is to initialize the cockroach swarm by using a uniform distribution of random integers for the initial dispersion pattern. After initializing the swarm, each of the operations for the cockroach behavior, or habit, executes -- chase, hunger, dispersion, and ruthlessness [14, 15]. The sigmoid function, see Equation 1, probabilistically transforms the position of each cockroach to move the colony through the search space.

Equation 1: Sigmoid Function

$$s(x_i) = \frac{1}{1 + exp^{-x_i}}$$

$$x_i = \begin{cases} 1 \; if \; random \leq sigmoid \\ 0 \; otherwise \end{cases}$$

The fitness value, or objective function, performs the move operation of the swarm toward a better or at least different global position, keeping the swarm from local minimum traps. This function is repeated until the end criteria is reached. In this case, the end criterion is the signal source.

The behavior and habits of cockroaches are based on the following formulation:

Table 2 - ICSO Equations [14, 15]

Chase:	$x_i = \begin{cases} w \cdot x_i + step \cdot rand \cdot (p_i - x_i), & x_i \neq p_i \\ w \cdot x_i + step \cdot rand \cdot (p_g - x_i), & x_i = p_i \end{cases}$	$i, g = 1, 2, \ldots, N.$
Ruthlessness:	$x_k = p_g$	$k \; is \; random \; integer$
Dispersion:	$x_i = x_i + rand(1, D),$	$i = 1,2, \ldots, N.$
Hunger:	$x_i = x_i + (x_i - ct) + x_{food}$	$i = 1,2, \ldots, N.$

The fitness function uses the second order differential to find changes in the search space. The search space is a hyperplane containing relatively flat geometry in the x-axis, frequency and y-axis, time. A deviation, or gradient, in the z-axis will indicate a signal. Since the background radio sources should be relatively stable with minor chirping, sensitivity provides escapement from local minimum traps or false alarms. The following equation is used as the fitness function:

$$f(x_{i,sample}) = ([expected \; frequency \; power] - [observed \; frequency \; power]) * (\frac{1}{tolerance})$$

5. Results

The four experiments with different settings of the ICSO variable set, Table 1, were completed. The variables under evaluation were the number of cockroaches (N), the hunger of the cockroaches (T), the step function (step), and the migration speed of the swarm (ct). The variables for the dimensional random vector (D=1,000) and inertial weight (w=0.1) were held constant. For example, Scenario 1, Table 3, consisted of 10 cockroaches that were not very hungry with a slow search speed. The scenario was run 10 times and averaged. This essentially baselines the false alarm rate as Type I and Type II Errors. Using 1,000 cockroaches decreased the false alarm rate and increased detection rates. The application of constant terms of the procedure were used to understand the sensitivity of the other variables with the success rate of detecting a signal, Table 4-6.

Table 3 - Scenario One

N=10; T=10; step=0.1; ct=0.2		
	Decision	
	Hit	Miss
Detect	4	1
No Detect	0	5

Table 4 - Scenario Two

N=100; T=100; step=1.1; ct=0.8		
	Decision	
	Hit	Miss
Detect	2	2
No Detect	0	6

Table 5 - Scenario Three

N=1,000; T=10; step=2.0; ct=0.5		
	Decision	
	Hit	Miss
Detect	5	2
No Detect	0	3

Table 6 - Scenario Four

N=1,000; T=100; step=0.2; ct=0.4		
	Decision	
	Hit	Miss
Detect	2	1
No Detect	0	7

Since the dimensionality of the data never changed, the *Chase* function controlled the speed of the swarm. The speed and step function sensitivity of the swarm helped better locate the minima within the data. The ICSO found signals with lower reporting of false negatives. Although the skewed results were likely due to the selection of the reduced data set, in all cases, the TYPE I Errors present indication that the underlying assumptions in variable selection and value sensitivities are not fully optimized.

6. Conclusion

Using meta-heuristic SI methodologies and other artificial intelligence techniques for radio astronomy data analysis is a relatively new venture in data mining. The strong case for using such techniques is the power in recognizing and classifying patterns with the data. The search for EC among the electromagnetic spectrum in the nearby H1 emissions bands continues unabated. Finding the perfect match of a meta-heuristic technique to radio signal data is an arduous task that takes much evaluation, planning, and experimentation, but using SI is useful in its ability to spread across the search space, collaboratively, in an attempt to find a maxima or minima pattern. Because the signal of an EC is a great unknown, the ferreting out of known signals in search of the unknown requires the use of elimination for negative detection. That is, topologies excluded within the search space are normal, and those without classification are of interest. If the EC exists, then identification and calibrated refinement of the SI search parameters should be able to help discover new and emerging patterns. Thus, swarm optimization techniques on radio telescope data with high dimensionality has strong potential [18].

References

[1] Oliver, Bernard M. 1994. "SETI [Search for Extra-Terrestrial Intelligence]." *IEEE Potentials*, August/September.

[2] *Proceedings of the IEEE.* 2012. "Beings on Earth: Is That All There Is?" May 13: 1700-1717.

[3] Phaidon Press Limited. 2017. *Universe.* London: Phaidon Press Limited.

[4] Hoskin, Michael. 1999. *The Cambridge Concise History of Astronomy.* Cambridge: Cambridge Press.

[5] SETI.org. 2017. *SETI Signal Searching.* December 3. Accessed December 3, 2017. https://www.seti.org/.

[6] —. 2017. *SETIQUEST DATA LINKS.* December 3. Accessed December 3, 2017. http://wiki.setiquest.info/index.php/SetiQuest_Data.

[7] IBM. 2017. *IBM Emerging Technologies Blog.* December 4. Accessed December 4, 2017. http://blog.ibmjstart.net/2016/06/29/ibm-stanford-university-team-new-perspective-seti-signal-analytics/.

[8] MIT. 2017. *Open Exoplanet Catalogue.* December 3. Accessed December 3, 2017. http://www.openexoplanetcatalogue.com/planet/Kepler-144%20b/.

[9] Smyth, John B. 1960. "Space Analysis of Radio Signals." JOURNAL OF RESEARCH of the National Bureau of Standards 293-297.

[10] Proesch, Roland. 2013. *Signal Analysis for Radio Monitoring.* Norderstedt: Books on Demand GmbH.

[11] *Proceedings of the IEEE.* 2009. "The Allen Telescope Array: The FirstWidefield, Panchromatic, Snapshot Radio Camera for Radio Astronomy and SETI." August: 1437-1447.

[12] Lockman, Felix J. 2017. "Radio Astronomy." *Observing the Invisible Universe.* The Smithsonian.

[13] Ross, Monte. 2006. "The New Search for E.T." *IEEE Spectrum* 32-37.

[14] Obagbuwa, I. C., and A. O. Adewumi. 2014. "An Improved Cockroach Swarm Optimization." *The Scientific World Journal* 1-14.

[15] Obagbuwa, Ibidun C, and Aderemi O. Adewumi. 2014. "An Improved Cockroach Swarm Optimization." *The Scientific World Journal* 1-13.

[16] Cullers, D. K, Ivan R Linscoft, and Bernard M Oliver. 1985. "Signal Processing in SETI." *ACM/IEEE-CS Joint Issue* 37-47.

[17] Gaber, Ali, R. Henry Tillman, R. Michael Buehrer, and Steve Ellingson. 2017. "Sensitivity Analysis of Localization Using Discrete Astronomical Radio Sources." *IEEE* 1-6.

[18] Korpela, Eric, Dan Werthimer, David Anderson, Jeff Cobb, and Matt Lebofsky. 2001. "SETI@HOME—MASSIVELY DISTRIBUTED COMPUTING FOR SETI." *Scientific Programming*, January/February: 78-83.

[19] ITNOW. “The Search for Little Green Men.” Oxford University Press. 2014

A Data-Driven Neuro-Wavelet Approach to Electric Arc Furnace Modeling

S. Ghaffaripour[a], K. Tirdad[a], A. Dela Cruz[a], H. Rahnama[b], A. Sadeghian[a]
[a]Dept. of Computer Science, Ryerson University, [b]RTA, Ryerson University

This paper proposes a hybrid data-driven neuro-wavelet based approach for modeling of the dynamic voltage-current (*v-i*) characteristic of the electric arc furnaces. The proposed modeling methodology is an integration of multi-scale filtering using wavelets, nonlinear regression techniques, and neural networks, and includes the following four main steps: data acquisition, preprocessing, feature extraction, and hybrid neuro-wavelet modeler. First, the voltage and current (*v-i*) information are collected from an operational electric arc furnace. Then, in the pre-processing step the measured *v-i* signals are transformed from time domain to time-frequency domain using a special form of wavelet decomposition analysis with the intention to gain a deeper insight into the data. In the feature extraction phase, the relevant wavelet coefficients that are more representative of the signals are identified. These coefficients are then used to form the modellers' input vector. In the final step, several artificial neural networks are built to find the functional relationships capable of mapping the arc current to voltage with the principal concern of minimizing error.

To assess the validity of the proposed method, four sets of experiments were conducted. These experiments examine the effect of changes in the number of decomposition levels of both voltage and current signals, and the choice of wavelet basis on the accuracy of the models. In order to investigate the accuracy of the neural network models, comparisons between actual data and the output of the neural networks were made (Figures 1 and 2).

We demonstrated that the wavelet transform lends itself particularly well to non-stationary signal analysis and modeling of electrical arc furnaces. It was shown that the time-scale representation of the voltage and current signals were capable of extracting more informative features in both time and frequency domains. Consequently, the learning procedure in the neural network mapping the current signal to the voltage signal became more effective and resulted both in fast convergence as well as negligible errors. It was specifically shown that neural networks, fed with such time-scale multi-resolution features and tuned with proper design parameters, can accurately model the underlying nonlinear/time-varying relationship between the arc furnace current and voltage after completing the learning process.

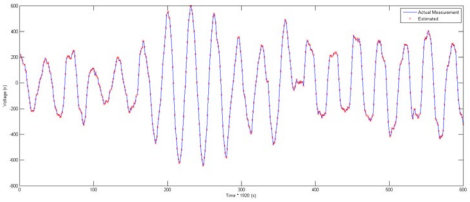


Fig. 1. The actual measured voltage vs. the estimated voltage

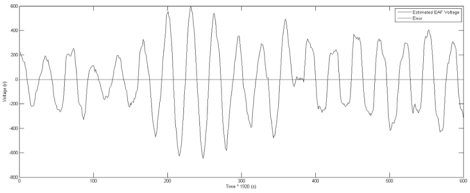


Fig. 2. The neuro-wavelet model response to the test data-case1

Part V

Thursday April 4th
11:10am-12:30pm
Contributed Talks 4:
Brain & Neural Systems
(Engineering Science 2008)

Dynamics of large amplitude fluctuations in human EEG differentiate individual and task-dependent variability

Kanika Bansal[1,2,*], Javier O. Garcia[1], Sarah F. Muldoon[3], Paul Sajda[2], and Jean M. Vettel[1,4,5]

[1]US Army Research Laboratory, Aberdeen Proving Ground, MD; [2]Department of Biomedical Engineering, Columbia University, New York, NY; [3]Mathematics Department, CDSE Program, and Neuroscience Program, University at Buffalo, SUNY, Buffalo, NY; [4]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA; [5]Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA
*phy.kanika@gmail.com

Scale-invariance has been observed in neuronal assemblies leading to the hypothesis that the human brain is functionally organized to operate near criticality [1]. Often, this criticality is demonstrated by the power-law distributions of neuronal avalanches, i.e., large bursts of neuronal activity, that are expected to display universal exponents and features. However, from the perspective of criticality, the dynamics of an active human brain during naturalistic cognitive processing is less understood [2]. Here, we investigated the bursts of large amplitude fluctuations in electroencephalogram (EEG) recordings of 36 healthy humans during rest and performance on two tasks which involved sequential viewing of emotionally charged images. Our analysis examined how the spatiotemporal distributions of these bursts or avalanches change during different tasks. We observed that the avalanche features change between tasks, indicating intra-subject dissimilarity across task states (Fig.1). This change is attributed to the localized shifts in avalanche probabilities during task performance. We also observed significant inter-subject variability in avalanche features for all the three conditions, indicating that the avalanches capture individualistic and cognitive features of the brain dynamics. Our results demonstrate the promise of this method for characterizing functional brain states and assessing individual variability.

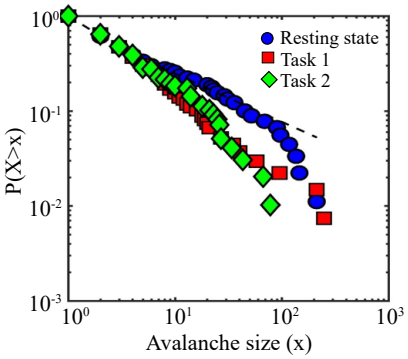


Figure 1: Probability distribution of avalanche size (x) within a subject for resting state and two different tasks. Resting state shows significant fit to a power law ($P(x) \sim x^{-\tau}$), with exponent $\tau = 1.9$. Dotted line shows the probability distribution for an ideal power law with $\tau = 1.9$. During task performance, avalanche features show deviation from the resting state distribution.

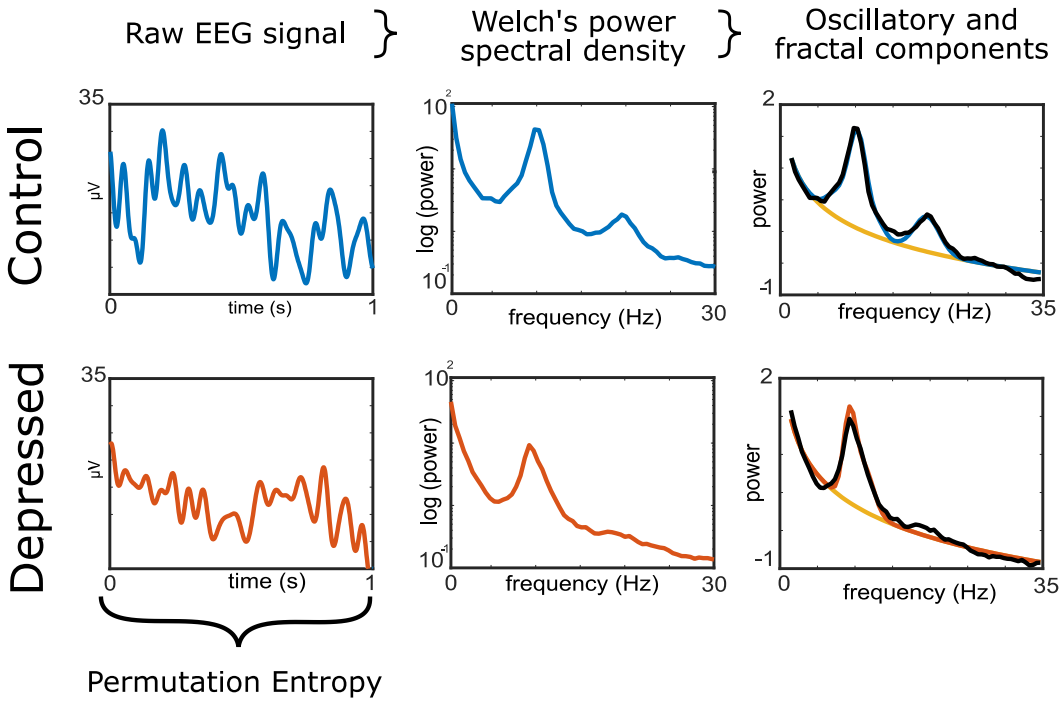
References

[1] Friedman et al. Phys. Rev. Lett. 108, 208102 (2012). [2] Shew et al. Nat. Phys. 11, 659663 (2015).

Modeling Electrocortical Power Spectra in Major Depressive Disorder

Maria Virginia Ruiz-Blondet, Carlos Martinez, Jourdan Pouliot and Vladimir Miskovic
Psychology Department, Binghamton University, State University of New York

The present work aims to explore whether there exist functional abnormalities of electroencephalographic (EEG) signals in major depression. We quantified the oscillatory and 1/f components of resting-state EEG recordings obtained from a public dataset [1]. The sample consist of 21 individuals with current or past episodes of major depressive disorder and 21 healthy controls. We used an automated parametric procedure to separately model the oscillatory and 1/f (fractal) components of the neural power spectrum [2]. Additionally, we also calculated the permutation entropy of time-domain EEG signals. We were not able to detect any reliable differences on any of the quantitative components of EEG signals between depressed individuals and healthy controls. We suggest that there are few obvious functional abnormalities of resting-state brain activity among young adults diagnosed with major depression. We hypothesize that previously reported differences are likely the outcome of heightened disease burden, in terms of substantial co-morbidity, greater chronicity, and severity of psychiatric symptoms.



References
[1] Cavanagh, J. F., Napolitano, A., Wu, C., & Mueen, A. (2017). The patient repository for EEG data+ computational tools (PRED+ CT). Frontiers in neuroinformatics, 11, 67.
[2] Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., ... & Voytek, B. (2018). Parameterizing neural power spectra. bioRxiv, 299859.

Changes in global brain connectivity resulting from Traumatic Brain Injury

Johan Nakuci[1], Mathew McGuire[1,2], Ferdinand Schweser[3], David Poulsen[2], Sarah F. Muldoon[1,4]

[1] Neuroscience Program, University at Buffalo, SUNY, Buffalo, NY
[2] Department of Neurosurgery, University at Buffalo, SUNY, Buffalo, NY
[3] Department of Neurology, University at Buffalo, SUNY, Buffalo, NY
[4] Department of Mathematics and CDSE Program, University at Buffalo, SUNY, Buffalo, NY

White matter tracts are important for efficient communication between distant brain regions and damage to these tracts can have significant impact on brain function. Traumatic Brain Injury (TBI) can cause axonal injury beyond the site of the trauma with white matter tracts exhibiting high vulnerability. Here we study changes in brain network structure resulting from injury using the later fluid-percussion model of TBI in rats. In injured rats, we find (1) a subnetwork of connections, distributed throughout the brain, with increased connectivity strengths when compared to control networks; (2) a shift in the network backbone toward the inclusion of shorter distances between regions when compared to controls and; (3) a subset of regions in which the prevalence of network motifs was altered. Furthermore, based upon individual differences in these structural changes across injured rats, we identify two subpopulations, reflecting different profiles of pathology induced changes. These results suggest that measures of global brain network structure can be used to quantify progressive changes in brain connectivity due to injury and to differentiate subpopulations of injury types.

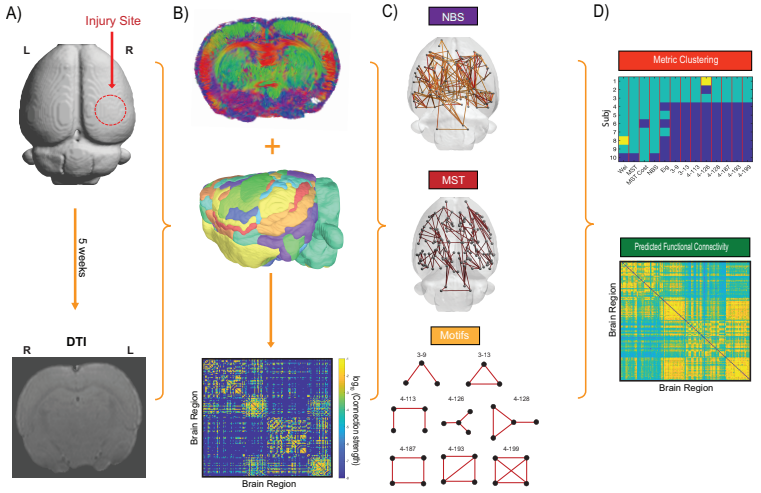


Figure 1. Analysis pipeline. **A**, Injury was induced in the right hemisphere using the lateral-fluid percussion model of TBI (Top). The red circle corresponds to the injury site. Five weeks after injury, diffusion tensor imaging was performed to assess brain structure (Bottom). **B**, Deterministic tractography (Top) was used to find connections between brain regions (Middle). An adjacency matrix was constructed based on the density of connections between regions (Bottom). **C**, Network based statistics (NBS; Top), minimum spanning tree (MST; Middle) and Motifs (Bottom) were calculated for injured and control brain networks. **D**, Clustering analysis was performed to identify subgroups within populations (Top), and predicted functional connectivity was modeled for these subgroups (Bottom).

Universality of the configuration-dynamics relationship in nonlinear networks.

Simone Evans, Anca Rădulescu

Mathematics, State University of New York, New Paltz

evanss3@hawkmail.newpaltz.edu

We study how architecture affects dynamics in nonlinear networks. First, we discuss our results from coupled quadratic nodes. While single-map complex quadratic iterations have been studied over the past century, considering ensembles of such functions, organized as coupled nodes in a network, generates new questions with potentially interesting applications to the life sciences. We discuss extensions of concepts like escape radius and Julia and Mandelbrot sets (as parameter loci in $\mathbb{C}^n$, where n is the size of the network).

One can conveniently use the system's asymptotic sets to calculate, visualize and interpret the long-term behavior of the system (in both phase and parameter spaces, with the network structure acting as a bifurcation parameter). The advantage is that of using clear topological markers (e.g., connectedness of a set) as the signature for the global dynamics of the system, amenable for prediction and classification. For example, we were able to identify an optimal parameter locus for which configurations which have identical Julia sets for one node-wise function (determined by the parameter c) do so for all c. This suggests means to produce robust classifications of ensemble behavior based entirely on the network architecture, independently on the node-wise dynamics (described by c). In this scenario, it is sufficient for the network to know in advance which hardwired structure is most effective to use in order to obtain a desired effect or avoid another, and then it can plastically modify its structure on a continuous basis, adapting online to new behavioral requirements.

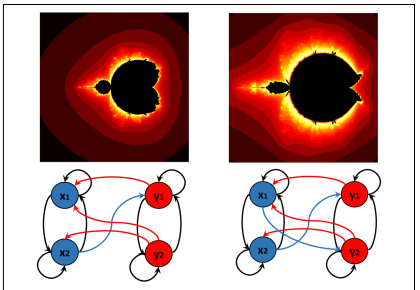


Figure: *Shape of the Mandelbrot set predicts symmetric versus non-symmetric network coupling.*

We then review ongoing research on two other nonlinear network models from neuroscience: threshold-linear networks and a reduced model of spiking inhibitory networks. Threshold-linear networks are networks that consist of simple, perceptron-like neurons with continuous-time dynamics. The inhibitory clusters model captures spiking activity and neuron synchronization in an all-to-all network of reduced Hodgkin-Huxley neurons. For each of these models, we seek out graph properties which can be used to predict or classify dynamics. Finally, we search for graph properties which are robust within each model, but that also translate between the three models.

Part VI

Thursday April 4th
3:45-5:05pm
Contributed Talks 5: Networks
(Symposium Hall)

Eigenvector-Based Centralities for Multilayer Networks are Tuned by the Topology of Interlayer Coupling

Dane Taylor[1]

[1] University at Buffalo, State University of New York (SUNY); danet@buffalo.edu

Centrality analysis refers to the study of node importances in complex networks, and it is widely used to understand the structural organization of complex systems. While there are numerous centrality analyses, it remains an active area of research to extend methods to more comprehensive network models including multilayer, multiplex and/or temporal networks. I will describe a centrality framework called *supracentrality* that extends eigenvector-based centralities (including pagerank, eigenvector centrality, and hub/authority scores) to multilayer, multiplex and temporal networks [1]. Given a network composed of layers $t = 1, 2, 3, \ldots$, our approach involves building centrality matrices $\{\mathbf{C}^{(t)}\}$ and coupling them into a "supracentrality matrix" $\mathbb{C}(\omega) = \text{diag}\left(\mathbf{C}^{(t)}\right) + \omega\tilde{\mathbf{A}} \otimes \mathbf{I}$. Here, $\otimes$ is the Kronecker product, $\omega \geq 0$ denotes the weight of interlayer coupling, and $\tilde{\mathbf{A}}$ denotes an interlayer adjacency matrix that gives the topology of interlayer coupling. Entries in the dominant eigenvector of $\mathbb{C}(\omega)$ give a centrality measure for each node i in each layer t, referred to as a *joint centrality*. We also define *marginal centrality* and *conditional centrality* to allow for a rich characterization of centrality. My talk will focus on new extensions to [1] exploring how *supracentralities* are influenced by the topology of interlayer coupling (i.e., by $\tilde{\mathbf{A}}$). I will describe mathematical findings using singular perturbation theory and empirical findings for systems including a multiplex airline transportation system for Europe and a temporal network encoding the graduation and hiring of Ph.D.s for U.S. Mathematics Departments.

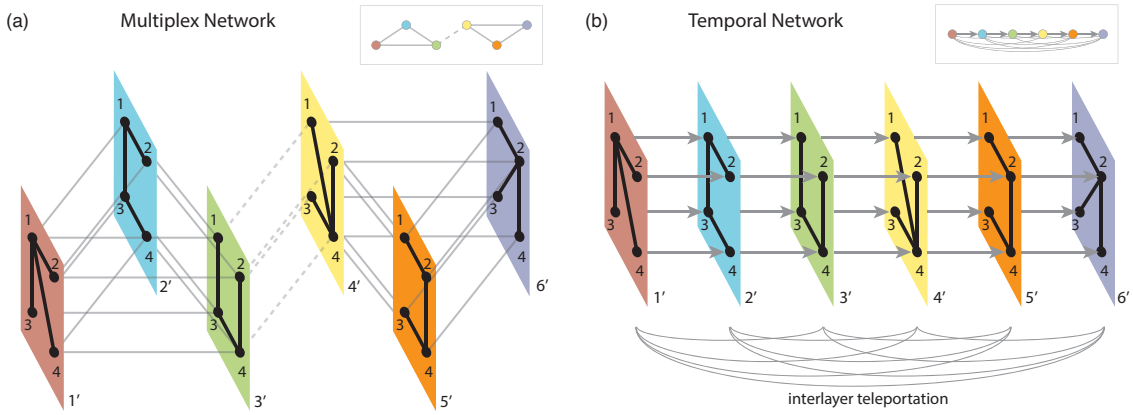


Figure 1: **Schematics of two types of multilayer networks.** (a) A multiplex network in which layers are coupled using an interlayer adjacency matrix $\tilde{\boldsymbol{A}}$ as illustrated in the inset. (b) A multiplex representation of a discrete-time temporal network, where we couple the sequence of layers through a directed (time-respecting) chain with interlayer teleportation. Formally, these types couplings for multiplex networks are called "diagonal" and "uniform." Note that our framework allows the interlayer couplings to be (a) undirected or (b) directed. In the directed case, node rankings are boosted for layers representing earlier times.

References

[1] D. Taylor, S. A. Meyers, A. Clauset, M. A. Porter and P. J. Mucha, *Multiscale Modeling and Simulation* 15(1), 537–574 (2017).

The Specialization Model for Network Growth

L. A. Bunimovich[1], D. Passey[2], D. Smith[2], and B. Webb[2],

[1] Georgia Institute of Technology
bunimovh@math.gatech.edu
[2] Brigham Young University
dj@math.byu.edu, dallas.smith@math.byu.edu, bwebb@math.byu.edu

Abstract

Dynamics in many real world complex networks exhibit a robustness to perturbations that is difficult to model. The difficulty arises when attempts are made to reproduce highly modular and hierarchical structure while also maintaining stability.

We introduce the specialization model of network growth to address this issue. As real world networks grow, sub-networks become specialized in preforming specific tasks. Our model mimics this behavior by maintaining and replicating the components of a network as it grows.

The specialization of a network via our process preserves spectral radius associated with the network's adjacency matrix. This allows us to show that a network maintains certain dynamic properties, specifically stability under mild conditions, as the network's topology becomes increasingly complex due to specialization.

Our specialized networks exhibit the desired increases in modularity, sparsity and hierarchical structure found in gene regulatory networks, the brain and the internet. Networks generated using our model also demonstrate many well known properties of real-wold networks such as the small-world property, disassortativity, power-law like degree distributions and clustering coefficients.

In addition to stability, specialization allows us to grow isolated components of the network. Using eigenvector centrality as a metric, we describe how the importance of these components evolves as they are specialized. Together these two ideas allow us to preserve the importance of some areas in the network while adjusting the importance of other areas in a structured way.

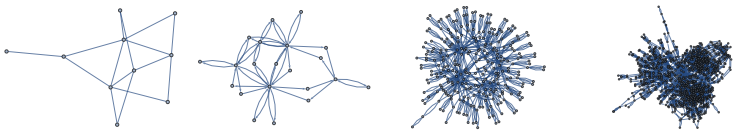


Figure 1: The stages of a network grown via the specialization model. The adjacency matrix of each network pictured has the same spectral radius.

Facility Location Using Markov Chains On Spatial Networks

Mitchell Sailsbery[1], Jacob Heiner[2], Connor Robertson[3], McKell Stauffer [4], and Tyler Jarvis[5]

[1] University at Buffalo, State University of New York
msailsbe@buffalo.edu
[2] University of Washington
[3] New Jersey Institute of Technology
[4] Brigham Young University
[5] Brigham Young University
jarvis@math.byu.edu

We model consumer-supplier interactions as a Markov chain on a spatial network. This model resembles the work by Faizrahnemoon et al. [1]. By using properties of this Markov chain, such as its Perron eigenvector, mean first passage times, and Kemeny constant, we develop several metrics for measuring optimal facility location. In several examples, we explore the interpretation of these metrics and compare their behavior to the p-median method, which is a more standard approach. These comparisons show that our network based metrics provide insight into the problem that is unobtainable by the p-median method, including redundancy of suppliers and the ability to maximize consumer throughput. We then present a case study where a religious organization builds another house of worship and is deciding where it should be located. We analyze publicly available data about church locations and interpret the results of both the network metrics and the p-median model. Finally, we discuss the limitations of this model as well as future developments, such as the development of algorithms to solve the formulated optimization problems as well as the scalability of the approach. Figure 1 shows an example consumer-supplier network where edges are weighted according to consumer traffic.

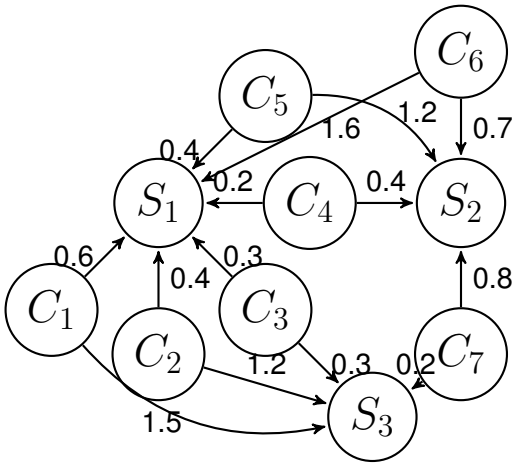


Figure 1: A sample graph including 7 consumers and 3 suppliers. Edge weights indicate distances, not probabilities.

References

[1] Mahsa Faizrahnemoon, Arieh Schlote, Lorenzo Maggi, Emanuele Crisostomi, and Robert Shorten (2015) A big-data model for multi-modal public transportation with application to macroscopic control and optimisation, International Journal of Control, 88:11, 2354-2368, DOI: 10.1080/00207179.2015.1043582

Spectrum Behavior of Laplacian for Multiplex Networks with General Coupling

Zhao Song[1] and Dane Taylor[2]

[1] Department of Mathematics
University at Buffalo, State University of New York
zhaosong@buffalo.edu
[2] Department of Mathematics
University at Buffalo, State University of New York
danet@buffalo.edu

Abstract

A leading approach for studying biological, technological, and social networks is to represent them by multiplex networks in which network layers encode different types of edges. Of particular interest in the study of Laplacian matrices for multiplex networks, called supra-Laplacians which arise for dynamics on networks, such as random walks and synchronization, and are also widely used for machine-learning algorithms. Our work builds on previous research analyzing the behavior of the spectrum of the supra-Laplacian matrix, which developed perturbation theory for the limits of strong and weak coupling between layers. Here, we extend this theory to characterize the eigenvectors and eigenvalues of supra-Laplacians for multiplex networks with general coupling between layers. In particular, we analyze multiplex networks with directed coupling between the layers, a situation that has not yet been explored.

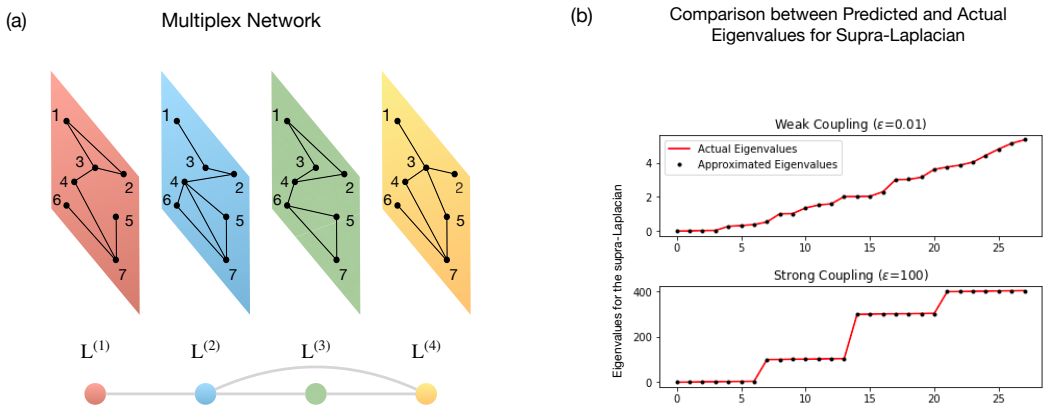


Figure 1: **Example multiplex network and its spectra.** (a) Multiplex network with $N = 7$ nodes and $T = 4$ layers. The layers are coupled using the topology shown underneath. (b) Predicted (symbols) and observed eigenvalues (lines) of the supra-Laplacian matrix for the limits of weak top) and strong interlayer coupling (bottom).

Part VII

Thursday April 4th
3:45-5:05pm
Contributed Talks 6:
Social Systems I
(Engineering Science 2008)

Dynamics of Urban Scaling

Lorraine Sugar, University of Toronto, Canada
Christopher Kennedy, University of Victoria, Canada

The observed phenomenon of urban scaling reflects a snapshot of an urban system in time. Scaling exponents are calculated from empirical data representing a single time period (i.e., a year) for all cities in a given country. In this research, we examine the dynamics of urban scaling by looking at scaling relationships over time, as well as in response to various urban development scenarios and policy mechanisms. We draw upon the *Origins of Scaling in Cities* model by Luis Bettencourt [1] to develop a new theoretical approach to understanding how scaling changes in different urban development scenarios. In particular, we focus on Bettencourt's new scaling parameter, G, which represents the city-system's balance between superlinear and sublinear scaling relationships.

In this paper, we provide a detailed overview of Bettencourt *Origins of Scaling in Cities* model, and then build upon the model to develop a new theoretical approach to the dynamics of urban scaling. We lay out a new theoretical derivation for changes in G over time that makes predictions consistent with the urban dynamics observed in empirical data for the U.S. over the period from 2004 to 2014. Specifically, we found that declining cities have a mean change in the parameter G that is statistically different, and consistently greater, than growing cities. In declining cities, G increases over time, which is also observable in the case studies of Pine Bluff, AK, and New Orleans, LA. The case study of Las Vegas, NV (Figure 1A), demonstrated a relatively consistent value of G, which is predicted by theory. We also modeled and tested the impacts of a few urban growth management policies on the change in G over time. For example, the case study of Lexington, KY (Figure 1B), demonstrated the drop in G over time for cities with urban containment policies, as predicted by our theory.

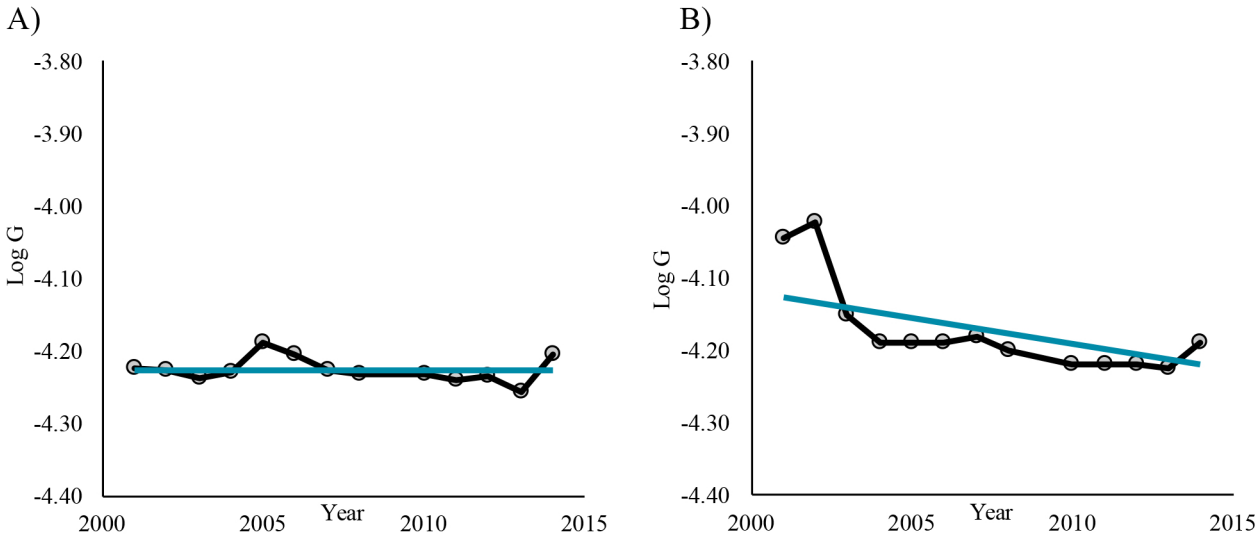


Figure 1: A) Rapidly growing city case study example of Las Vegas, NV; B) Urban containment policy city case study example of Lexington, KY.

References

1. Bettencourt, L. 2013. "The Origins of Scaling in Cities." *Science* 340(6139): 1438-41.

CDS Central Counterparty Clearing Liquidation: Road to Recovery or Invitation to Predation?

Magdalena Tywoniuk*†‡

Department of Finance, University of Geneva & Swiss Finance Institute

April 2017

Abstract

Recent regulation, mandating the clearing of credit default swaps (CDS) by a Central Clearing Counterparties (CCP), has rendered it's possible failure a serious threat to global financial stability. This work investigates the potential failure of a CCP initiated by the default of a large dealer bank and the unwinding of its positions. The theoretical model examines variation margin exchange between dealer banks and the price impact of liquidation and predatory selling. It provides a measure of covariance between assets in banks' portfolios; price impact affects assets to varying degrees, based on their relative distance to defaulted assets. Key results show that liquidation lowers CCP profits, and how predation decreases the profits of all members, pushing banks to default. Furthermore, a hybrid CCP (vs. current) structure provides a natural disciplinary mechanism for predation. Also, it is more incentive compatible for the CCP, in expectation of a large loss. A multi-period, dynamic simulation, calibrated to OTC market data, provides parameter sensitivities concerning the magnitude of CCP and predatory bank gains/losses, specifically, the minimisation of those losses with a hybrid fund structure. Furthermore, regulatory implications concerning the timing of liquidity injection for a Lender of Last Resort (LoL) are determined for various liquidity scenarios; stable and decreasing market liquidity, as well as, a liquidity dry-up at the bottom of a financial crisis.

Keywords: Systemic Risk, CCP Recovery, CDS, CDS Spread Fire Sales, Liquidation, Predation, Price Impact, Contagion, Financial Network, Over the Counter Markets.

JEL Classification: G00, G01, G02, G14, G10, G18, G20, G23, G33

*Electronic address: Magdalena.Tywoniuk@unige.ch

†Av. de Bethusy 26, 1005, Lausanne

‡+41 788255110

7.1 Under normal market liquidity

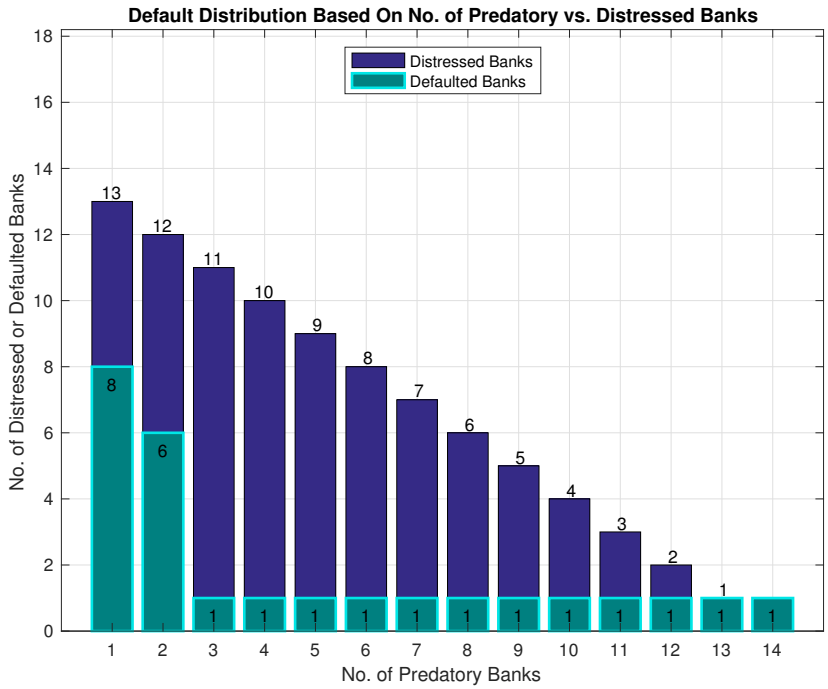


Figure 2: Number of distressed banks as drivers of defaults (no collusion, increasing predators).

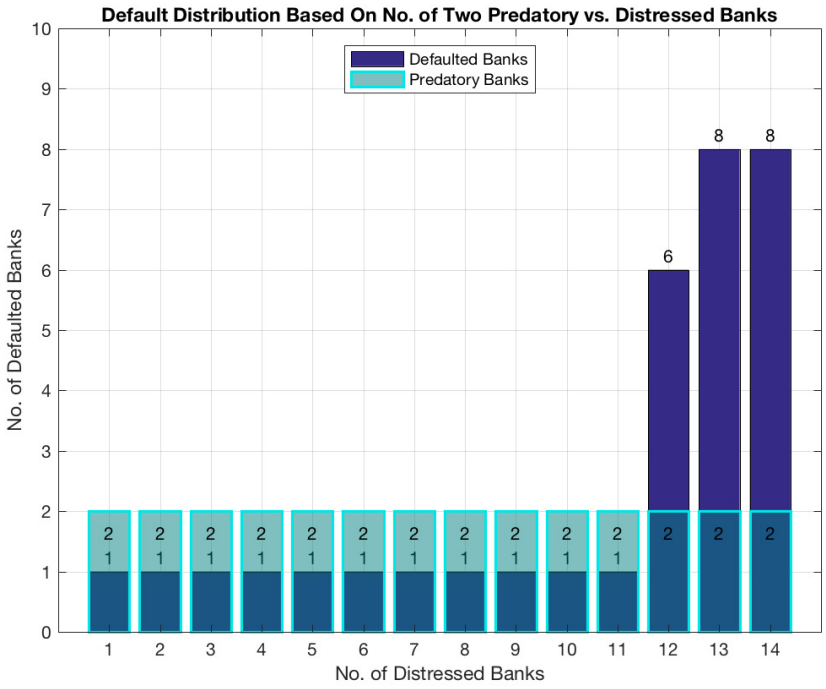


Figure 3: Higher defaults driven by increasing distressed banks vs. predator competition.

The Emergence of Monitoring

Andreas D Pape[1] and Peter DiCola[2]

[1] Department of Economics &
Center for Collective Dynamics of Complex Systems
Binghamton University, State University of New York
apape@binghamton.edu
[2] Pritzker School of Law
Northwestern University
p-dicola@law.northwestern.edu

Ostrom's famous Principles of Polycentric Governance[1] are characteristics of institutions that manage common pool resources (CPRs) that she found, from case studies, tend to describe those that successfully administer those CPRs. This prompts the observation: If these principles tend to improve management of CPRs by a community, and successful management would increase the long-term viability of the community, then in an appropriately specified evolutionary, agent-based model, we should see the emergence of these principles. This project attempts to validate this observation.

One principle is monitoring, which is that agents in the community can and do keep track of whether others in the community follow the rules that the community has chosen to manage the CPR. One issue of importance is how to evaluate the emergence of an institution; we attempt to define the agents' behavior broadly so that some behaviors (closely following other fishermen on a lake) can be interpreted ex-post as monitoring without an ex-ante specification (i.e. agents simply choose "monitoring" as a pre-specified action available.)

References

[1] Ostrom, E., 2015. *Governing the commons.* Cambridge university press.

Iterated Function Analysis of Social Convention Changing Process

Manqing Ma(mam6@rpi.edu)[1] and Jianxi Gao(gaoj8@rpi.edu)[1]

[1] Rensselaer Polytechnic Institute, Troy, New York

The "tipping point", i.e the "critical mass" in the social convention changing process, is the minimum percentage of committed users of the opposite opinion aiming to convert the whole society. In [1], the authors have found this to be around 0.25 for a complete network in both the artificial setting and the real-world experiment setting, in which people's average memory length is around 12. In this paper, we aimed to further explore the theoretical aspects of this result. We proposed a iterated function for the original agent-based modeling. We created a mapping between the "memory length" defined in the original experiment setting and the nodes' average in-degree in our modeling. We argue that the original simulation is in fact with multiple fixing points, which means that the network will always come to a stable state with everyone keeping their opinions unchanged, regardless of their initial opinions, or the percentage of the committed users in the system. The convergence, and the status of the network at convergence are discussed in the paper. Using this model one could look further into how memory length could influence the "tipping point", by studying the convergence of the iterated function.

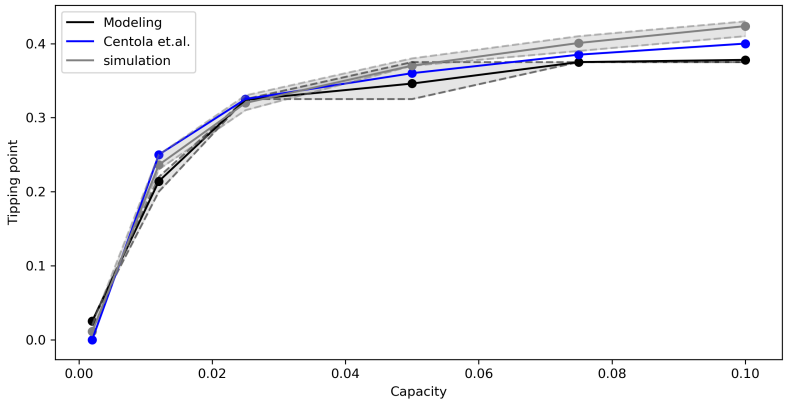


Figure 1: Comparison between the model(regular graph, capacity:linking probability), simulation(complete graph, capacity:Memory length) and Centola et.al.'s result (capacity:Memory length)

References

[1] Centola, Damon, et al. "Experimental evidence for tipping points in social convention." Science 360.6393 (2018): 1116-1119.

Part VIII

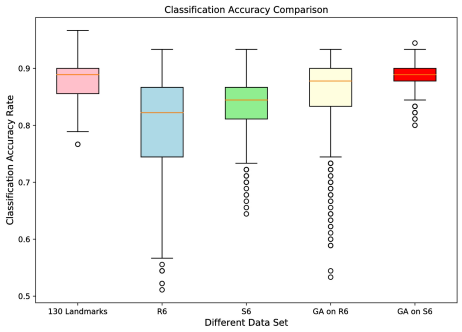
Friday April 5th
11:10am-12:30pm
Contributed Talks 7:
Data Science Special Session II
(Symposium Hall)

Optimizing Facial Feature Extraction for Emotion Detection on Mobile Devices

Yingjun Dong* and Hiroki Sayama
Department of Systems Science and Industrial Engineering, Binghamton University
* ydong25@binghamton.edu

Light-weight devices such as smartphones are used widely in daily human life. Studies have been conducted on how to implement human-computer interaction (HCI) on light-weight devices. Facial expression recognition (FER) is one of essential tasks in the HCI domain. However, the commonly used methods for FER are computationally demanding, using 68~130 landmarks [1,2]. In this study, to achieve similar FER efficiently on light-weight devices, we aimed at decreasing the computational overhead by reducing the number of landmarks. We used the Manual Annotation on AR Face Database [2,3] which contains 130 facial landmarks in 4 emotion expressions for 112 human subjects. To select fewer facial landmarks, we first calculated mutual information (MI) between every pair of landmarks in terms of their displacements across different expressions, which were summarized in the form of a distance matrix. Then we applied hierarchical clustering to the MI distance matrix. As a result, we obtained 6 clusters of facial landmarks (Figure 1a). To simplify the computation, we randomly selected one facial landmark from each cluster or six landmarks from the whole 130 landmarks, and used only those 6 landmarks for classification. We used a Support Vector Machine (SVM) classifier to conduct classification of facial expressions on Raspberry Pi (used as a substitute platform of lightweight mobile devices) using either the original 130 landmarks or the 6 selected landmarks. We then applied the genetic algorithm (GA) to optimize the selection of 6 landmarks, either selecting one from each cluster or selecting six from the whole 130 landmarks. We compared the classification accuracy of 130 landmarks, random select 6 landmarks (R6), selected 6 landmarks (S6), applied GA on R6 and applied GA on S6 (Figure 1b). Results (Figure 1c) showed that we successfully reduced the computational time by 60% while keeping the classification accuracy at about the same level. It was also found that the facial landmark clustering significantly helped the effective landmark selection.

(a) (b) (c)



Computational Time Comparison	
Selected 6 Landmarks	54.01 seconds
130 Landmarks	148.16 seconds

Figure 1: (a) Clustered facial landmarks (clusters shown in colors). (b) Classification accuracy results. Each condition was repeated 1,000 times. (c) Computational time comparison.

[1] Gross, R., *et al.* (2008). *Proc. 8th IEEE Intl. Conf. Automatic Face and Gesture Recognition*.
[2] Martinez, A., & Benavente, R. (1998). *Computer Vision Center Technical Report* 24.
[3] Ding, L., & Martinez, A. M. (2010). *IEEE Trans. Pattern Anal. Mach. Intell. 32*(11).

Multi-fidelity Modeling in Sequential Design for Identification of Stability Region in Dynamic Time-Delay Systems

Yiming Che, Changqing Cheng[1]
Department of Systems Science and Industrial Engineering
State University of New York at Binghamton, Binghamton NY 13902

Abstract: Time delay arises in a variety of real-world complex systems. High-fidelity simulation generally renders high accuracy to simulate dynamic evolution of such complex systems and appraise quantity of interest for process design and response optimization. Identification of limit states exemplifies such applications, which explicates the boundary that separates distinct regions (e.g., stability region) in design space. While design of experiments is the common procedure to evaluate decision functions to sketch the boundary, it crucially relies on the quantity and quality of sampling points. This has made it infeasible to explore large design spaces with expensive-to-evaluate high-fidelity simulations. Furthermore, the complicated contour of stability region in time-delay systems nullifies most existing sequential design paradigms, including adaptive classification approaches. On the other hand, low-fidelity surrogate modeling efficiently emulates high-fidelity simulation, albeit at the expense of accuracy, not ideal to inspect system behavior near the boundary. We investigate a multi-fidelity approach to spot stability region in a sequential fashion: high-fidelity simulation is only called upon to evaluate sampling points selected conforming to the exploration-exploitation principle, and surrogate modeling is then deployed to approximate the decision function.

The flowchart of the proposed approach is displayed in Figure 1: a space filling algorithm Latin hypercube design (LHD) is first used to initialize the sampling points across the design space. Quantity of interest at those design points is evaluated by high-fidelity simulation (e.g., finite element method), which is then assimilated into the training set D for the learning of low-fidelity surrogate. In our study, Gaussian process (GP) is used as low-fidelity surrogate. Sliced Latin hypercube design (SLHD) is utilized to generate next batch (slice) of sampling points: design points on each slice as well as the collection over all the slices still form a LHD. Thus, the whole design space can be more effectively explored. Afterwards, surrogate modeling is applied to those new sampling points from SLHD slice by slice to evaluate the mean and variance of the quantity of interest and hence approximate the decision boundary. Subsequently, critical points among those samples from the new slice are handpicked, conforming to the expected improvement (EI) criterion that the critical point will appear at the largest EI value, to be further assessed via high-fidelity simulation. Here, EI is designed to automatically balance the tradeoff between exploration and exploitation for the selection of critical points. As this process iterates, the number of selected critical points in training set D will increase and thus the estimated boundary converges gradually. Pivotal to our algorithm is that the surrogate does not have to be globally accurate, as only the vicinity of the target boundary is critical. Consequently, the estimated boundary can be obtained by much less function evaluations compared to one–shot design.

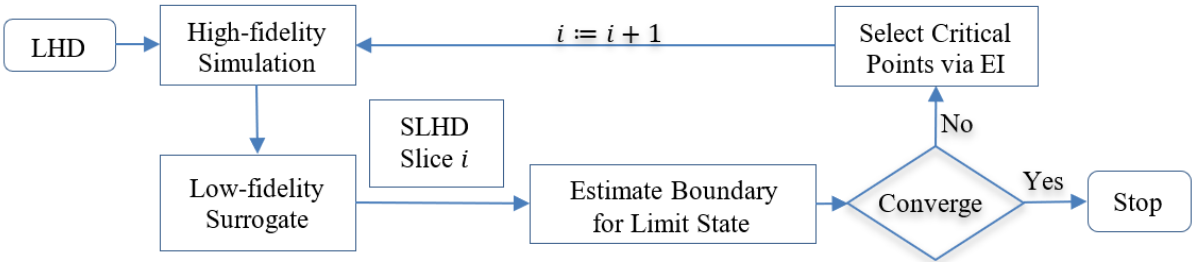


Figure 1 Flowchart of Bayesian optimization based sequential design

The application in the numerical case study of delayed Mathieu equation with disjoint stability region as well as real world machining process with continuous stability region corroborates the proposed approach. The boundaries that separates stability region in both cases are successfully estimated by the proposed method. For the first case, only 325 training points are used to estimate the boundary obtained by 10000 points in one-shot design and for the second case, 400 training points are used to obtain the estimated boundary simulated by 10000 points in one-shot design

Keywords: Multi-fidelity approach, limit state, sequential design, time-delay system

[1] Corresponding author: ccheng@binghamton.edu

Efficient algorithms for crowdsourcing problems introduce bias

Abigail Hotaling and James P. Bagrow University of Vermont, USA

Crowdsourcing is a modern technique to solve computationally challenging and complex sets of problems using the power of human participants [1]. However, human participants are relatively expensive compared with computational methods, so considerable research has investigated algorithmic strategies for efficiently distributing problems to participants and determining when problems have been sufficiently completed [2, 3]. Yet the effects of these algorithms are not well understood, in particular which problems are sufficiently answered and which problems remain unanswered when using these algorithms. Here we show that algorithms that efficiently solve crowdsourcing problems do so by introducing bias against difficult problems. This bias leads to an over-representation of easy problems in the set of completed problems and a corresponding under-representation of hard problems (Fig. 1).

To study crowdsourcing algorithms and problem bias we introduce a model for crowdsourcing a set of problems where we can tune the distribution of problem difficulty. If an unbiased method could be used, then the distribution of difficulties for its completed problems would be the same as the distribution for all problems. We then apply a crowdsourcing algorithm, *Requallo* [2], to our model and find that the set of problems it completed is biased (Fig. 1). Finally, we introduce an inverse procedure to estimate the true problem difficulty distribution given only an algorithm's responses, allowing us to reason about the larger problem space with limited information.

One of the most common applications of crowdsourcing is gathering training data for machine learning algorithms. A consequence of efficient method-induced bias is that training data derived from efficient crowdsourcing will lack labels for difficult problems. Yet those labels are often the most important information to provide for learning. Therefore, while it is more costly to forgo efficient crowdsourcing methods, minimizing bias may be worth the added expense to maximize machine learning accuracy.

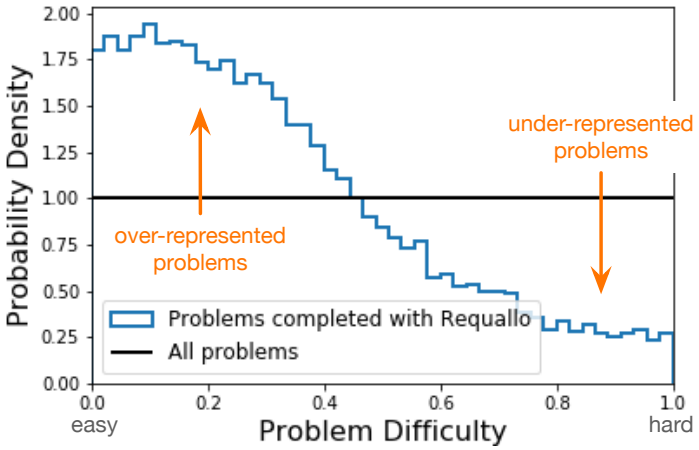


Figure 1: Efficient algorithms introduce bias. An unbiased method would recover the original uniform distribution of problem difficulty among completed tasks. However, the efficient Requallo method [2] over-represents easy problems and under-represents hard problems.

[1] Brabham, D. C. (2008). Convergence, 14(1), 75-90.
[2] Li, Q., Ma, F., Gao, J., Su, L., & Quinn, C. J. (2016). In WSDM'16, ACM.
[3] McAndrew, T. C., Guseva, E. A., & Bagrow, J. P. (2017). PloS one, 12(8), e0182662.

UAFS: Uncertainty Aware Feature Selection for Missing Data Problems

Andrew Becker[1,2] and James Bagrow[1,3]

[1] Mathematics & Statistics and Vermont Complex Systems Center,
University of Vermont
[2] ajbecker@uvm.edu [3] james.bagrow@uvm.edu

Missing observations and incomplete data are common in real world data sets [1]. Imputation methods [2], that attempt to infer the values of unseen variables, are crucial for dealing with missing data when performing statistical analysis and supervised learning [3]. Yet excessive missingness and high dimensionality in data sets that are being imputed challenges most methods. Here we introduce a method, *Uncertainty Aware Feature Selection* (UAFS), that handles extreme levels of missingness in imputation problems. UAFS works by selecting subsets of variables that appear most informative or most significantly related to an outcome variable we wish to predict while accounting for the missing data in those variables. Combining UAFS with imputation leads to improved performance across many problems and data sets: for example in Fig. 1, UAFS achieves the same imputation accuracy at 75% missingness as occurs at 50% missingness without UAFS. UAFS results in straightforward and objective subset selection for imputation problems and works across a variety of data types and has extensions into non-supervised learning problems. In general, improved performance in imputation results in improved predictive accuracy when building models, thus allowing machine learning to push into data-driven studies of complex systems where missingness was previously problematic.

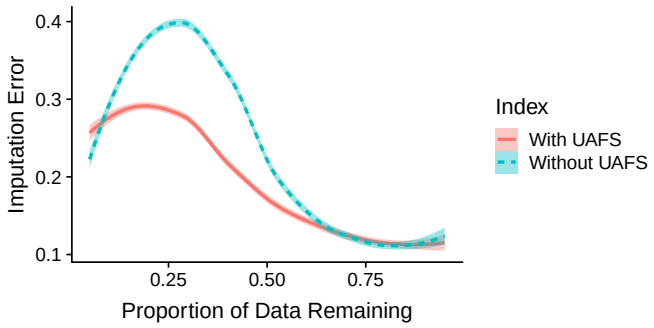


Figure 1: **Imputation Accuracy with UAFS**. UAFS significantly reduces imputation error (RMSE), especially at high values of missingness.

References

[1] J. Scheffer, "Dealing with missing data," *Research Letters in the Information and Mathematical Sciences*, vol. 3, pp. 153–160, 2002.

[2] R. J. Little and D. B. Rubin, *Statistical analysis with missing data*. John Wiley & Sons, 2014.

[3] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013.

Part IX

Friday April 5th
11:10am-12:30pm
Contributed Talks 8:
Social Systems II
(Engineering Science 2008)

Modeling the Cooperative and Adversarial Behaviors of Farmer and Regulator Agents in Vermont's Missisquoi Bay Area

Kevin Andrew[1]

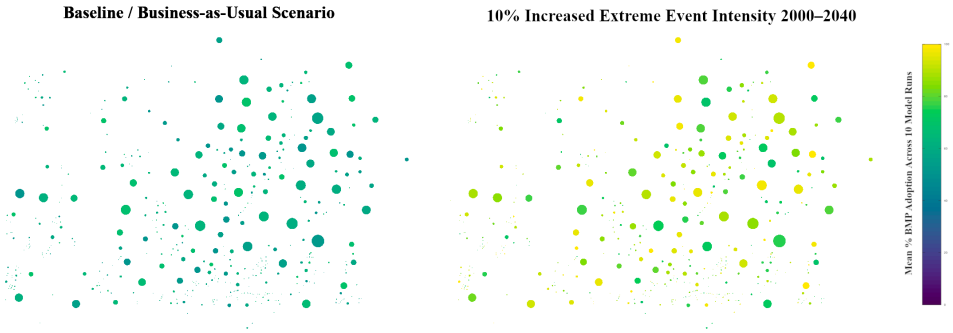
[1] University of Vermont
kandrew6@uvm.edu

Projected climatological changes over the next century are expected to drastically alter the weather patterns of the Northeastern United States. Combined with anthropogenic changes to land use and projected shifts in population distribution and urban development, the future health of aquatic environments within the region remains uncertain.

As part of the development of a comprehensive integrated assessment model (IAM) for the Lake Champlain Basin, the behavior of farmer and governmental regulator agents in the Missisquoi Bay Area of Lake Champlain in Vermont was modeled to explore potential changes in human decision-making and land-use under a variety of projected climatological, environmental, and economical scenarios for the region to the year 2040. In particular, we are looking at how farmers within the region may choose to change their land-use practices and adopt or reject agricultural best management practices (BMPs) and how a government regulator may implement taxes on or subsidize farming practices in an attempt to stymie environmental damages to the lake ecosystem.

480 farmer agents, corresponding to the agricultural land parcels within the Missisquoi Bay Area, and 1 municipal regulatory agent were included in the model. The behavior of these agents was trained under 14 scenarios, 13 projected and 1 baseline/business-as-usual, using deep reinforcement machine learning with double q-learning (DDQN). A comparative analysis of the results of the model under these various scenarios will be discussed, such as those seen in Figure 1, comparing the mean BMP adoption of farmer agents in the model across 10 runs, trained for both the baseline scenario and a scenario with a 10% increase in extreme weather event intensity.

Figure 1: Modeled Changes in Mean Adoption of BMPs of Farmer Agents in the Bay Area between a Baseline Scenario (Left) and a Scenario with a 10% Increase in Extreme Weather Event Intensity (Right) Across 10 Model Runs. Point Size and Relative Location Corresponding to Size and Distribution of Underlying Land Parcels.



Collective phenomena in sensorimotor synchronization: How the collective central moment stabilizes group drumming

Dobromir Dotov[1] and Laurel Trainor[1,2]

[1] LIVELab, PNB, McMaster University and [2] Rotman Institute, Ontario
dotovd@mcmaster.ca, ljt@mcmaster.ca

Collective phenomena make for an important topic in the study of animal behavior, large scale human collectives. Less is known about how small groups of humans coordinate their movements to achieve a shared goal, in particular when the nature of the task is sensorimotor. We considered group drumming as a relevant and feasible paradigm.

Groups of four participants on bucket drums maintained a steady tempo and synchronized with each other. All trials were repeated in a solo condition. Performance in the group condition was analysed as Individuals-in-Group (individual participants' timing data) and Group, where the group-level drum hits were defined as the central moments of the individual drum hits. Thus, the following three conditions were compared: Solo, Individuals-in-Group, and Group.

Each time-series of inter-beat intervals obtained from the drum hits in a trial was analysed using, 1) lagged auto- and cross-correlations and, 2) the parameters of a drift-diffusion model. First, negative auto-correlations in Group, Solo, and Individuals-in-Group indicated a self-correcting dynamic but, importantly, the positive cross-correlations at both lag-0 and lag-1 implied mutual anticipation. Second, the attractor strength in Group was highest, followed by Solo, and Individuals-in-Group. Noise was lower in Group than in Solo and Individuals-in-Group, see Figure 1.

The fact that collective performance stabilized group-level behavior and enabled the emergence of mutual anticipation could be explained in terms of the central moment of group performance, consistent with models of group animal motion [1], and akin to a mean field created by the sum of individual units' movement.

Such studies address the collective origin of music. Music is a universal and very early human phenomenon that binds private experiences and collectively coordinated sensorimotor behaviors.

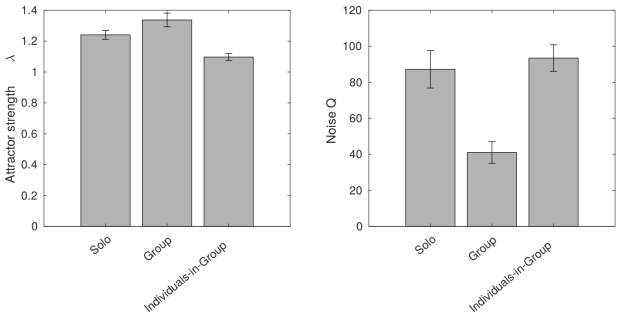


Figure 1: Stability and noise in the three collective modalities.

References

[1] Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1); Sumpter, D. J. T. (2006). The principles of collective animal behaviour. *Philosophical Transactions of the Royal Society B*, 361(1465).

Information diffusion in organizations: A network simulation

Jason Marshall, Neil G. MacLaren, Siaki Tetteh-Nartey, and Etka Topaloglu
Binghamton University, State University of New York, jmarsha7@binghamton.edu

Identifying the parameters that optimize information diffusion and quality retention could have profound implications for organizations. Previous work has attempted to identify ideal network parameter values for idea diffusion using a trade-off optimization approach. Following these authors, we used an SI model framework to compare a benefit/cost ratio (BCR) of information diffusion on randomly generated realistic networks. We developed random edge-colored graphs with two edge types representing official or unofficial social ties; an information vector transferred more easily along the matching edge type in an asynchoronous update regime. Unofficial edges were generated with a truncated power law degree distribution. Organization edges were added independently to create a connected, branching, locally clustered graph with random weak ties. Nodes had random listening, transmitting, and switching parameters that determined the probability a node would receive and transfer information along the same or opposite edge type. Holding graph size constant ($\nu = 100$), the maximum number of official edges, maximum unofficial edge degree, slope of the power law unofficial degree distribution, network switching probability, and originator closeness centrality were varied over 9,600 simulation runs. A decay constant, held equal for the two information types to avoid biasing results, reduced the information quality at each new node receiving the information. We counted the number of nodes that received the information vector after a fixed number of steps; we also recorded the graph average information quality among nodes that received the information and the number of times the information vector switched edge types. Many simulation runs did not show any information movement, so linear regression analyses were conducted on 4,137 runs where more than one node, besides the originator, heard the information. Network topology parameters generally had statistically noticeable but practically negligable effects. Unofficial information tended to reach more nodes for a given level of decay and larger dominant eigenvalues of the graph adjacency spectrum were associated with information reaching fewer nodes for a given level of decay. We found that a high number of switches and high originator closeness centrality maximized the BCR.

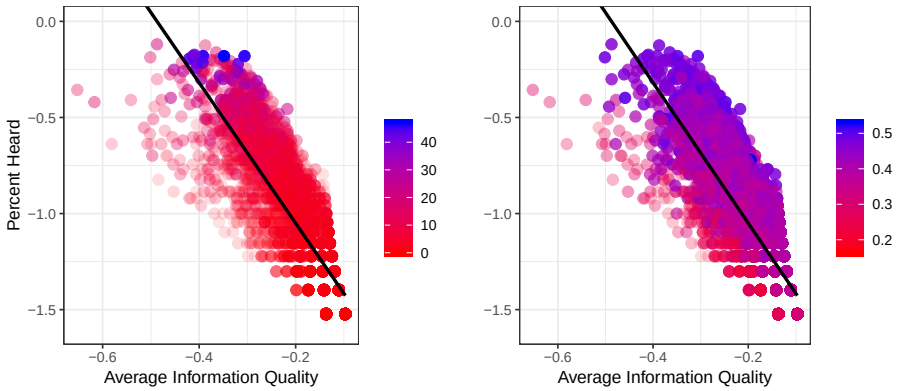


Figure 1: Simulation runs showing the cost (average information quality) and benefit (percent heard) observables on a log-log scale with colors representing the number of network switches (left) and originator closeness centrality (right). Regression line is $y = -1.98 - 4.53x$ with $R^2 = 0.93$.

Examining the Effects of Expertise Diversity on Collective Design and Innovation Using an Online Social Network Experiment and "Idea Geography" Visualization: An Initial Report

Y Cao, Y Dong, M Kim, NG MacLaren, A Kulkarni, SD Dionne, FJ Yammarino, & H Sayama
Binghamton University, Binghamton, NY 13902, USA. ycao20@binghamton.edu

This study aims to experimentally investigate how organizational network structure and diverse expertise of participants affect collective design and innovation. We recruited a multidisciplinary group of students at a mid-size public university in Fall 2018 to participate in an online collective design experiment for two weeks. Participants were initially asked to provide a written description of their major, career interest and expertise. We used Doc2Vec to convert these narrative data to 400-dimensional expertise vectors. Using this information, participants were assigned to three groups with different expertise distribution: (1) spatially clustered, (2) randomly distributed, and (3) dispersed. The network topology was a spatially clustered network with average degree four in all three conditions. On a custom-made web-based platform with an interface similar to Twitter, the participants collaboratively worked with their social neighbors to create a catch phrase for marketing a laptop. The daily ideas produced by them were recorded electronically on each working day during the two-week session. After the experiment, the participants were asked to submit their favorite final designs, which were then evaluated by experts (Marketing PhD students who did not participate in the experiment) on a 5-point Likert scale. We converted the daily ideas and final designs to 400-dimensional vectors again using Doc2Vec. Principal component analysis was applied to the whole collection of idea vectors for dimensionality reduction and visualization (Fig. 1a). The variance of idea distances from the "center of mass" is shown in Fig. 1b. The average evaluation score of each final idea was used as its elevation to reconstruct utility terrains for each of three groups in the 2-D principal component space, which we call "idea geography" (Fig. 1c). We found that there were clearly identifiable utility mountain areas in each of three idea geography spaces, where most of the submitted final designs were concentrated. Group 2 (with randomly distributed expertise) came up with the best final ideas among the three groups in terms of idea utility values. Group 3 (with dispersed expertise) came up with two optimal ideas that were next to the best idea of Group 2. In contrast, Group 1 (with clustered expertise) generated many optimal final ideas, which are seen as multiple peaks in their idea geography. We plan to continue to conduct this online experiment to collect more data and conduct more systematic analyses. This material is based upon work supported by the National Science Foundation under Grant #1734147.

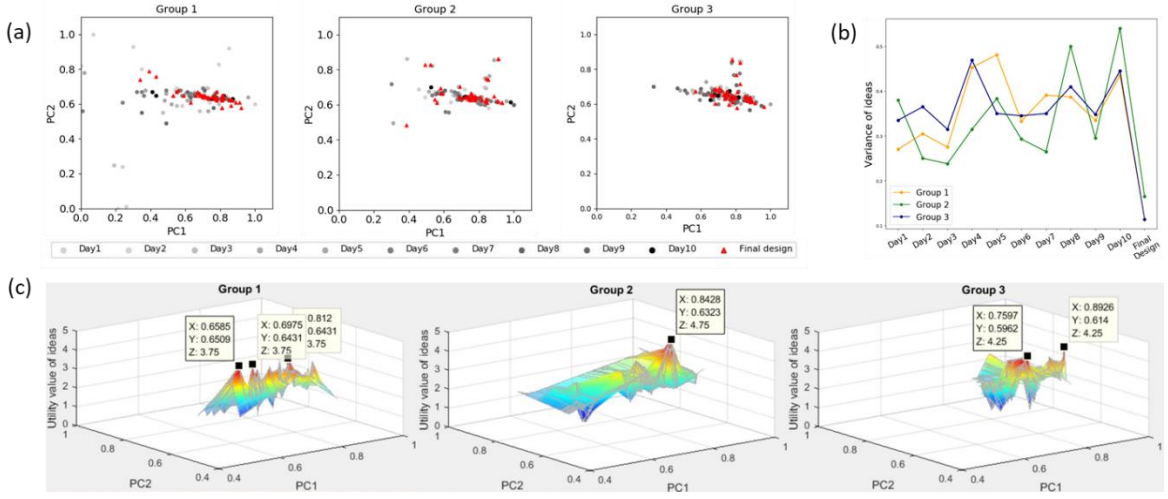


Figure 1: (a) Distribution of daily and final ideas. (b) Variance of daily and final ideas. (c) 3-D surface plot of "idea geography" constructed using evaluations of final designs.

Part X
Posters

Identification of Early-Stage Atrial Fibrillation with Intrinsic Properties of RR intervals

Yiming Che, Changqing Cheng[1]
Department of Systems Science and Industrial Engineering
State University of New York at Binghamton, Binghamton NY 13902

Abstract: Atrial fibrillation (AF) is a common cardiac arrhythmia and is responsible for a number of complications. While early-stage AF typically lasts only a few episodes and may not be immediately life-threatening, the cardiac arrhythmia favors electrical and structural alteration of the atria that tends to intensify and perpetuate AF even at incipient stage. Therefore, early-stage identification of AF episodes is required for effective treatment and disease management. Nonetheless, the brevity of early-stage AF negates a myriad of conventional models for effective detection. In this paper, we propose novel biomarkers to quantify energy and its entropy based upon intrinsic time-scale decomposition (ITD) to parse single-lead ECG signals collected from wearable devices. ITD provides accurate instantaneous time-frequency-energy charactersitics for the nonlinear and nonstationary data, particularly the short-term time series signals. Using datasets from physionet.org, our model registers average accuracy of 95%, specificity of 96% and sensitivity of 93% for the diagnosis of AF events, handily outperforming wavelet-based algorithm.

In this present study, we investigate the instantaneous TFE characteristics of RR intervals to differentiate AF and SR signals. Again, short-term single-lead ECG signals collected from portable ECG device AliveCor are used. Therefore, this study is amenable for out-of-hospital monitoring and early detection of AF. The data are retrieved from physionet.org, where ECG signals with duration between 9 seconds and slightly above 1 minute are provided, each of which is annotated as AF or SR by domain experts. The proposed ITD-based algorithm is summarized in Fig. 1.

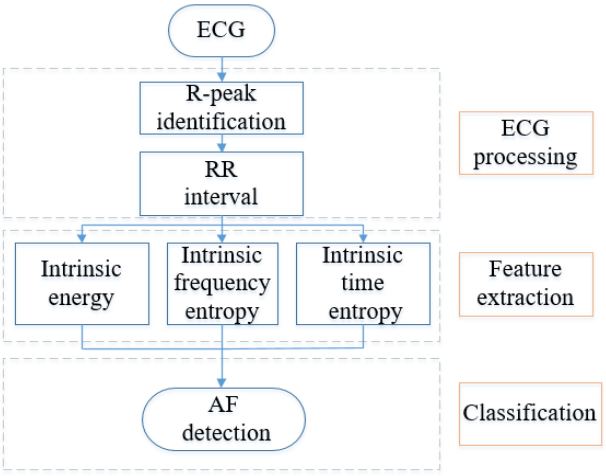


Fig. 1 Flowchart of the proposed algorithm

The well-known Pan-Tompkins algorithm is first adopted to identify R-peaks from the raw ECG signals, then the time series of RR intervals are extracted. It bears mentioning that it is fairly onerous to directly tie the raw ECG signals with the status of cardiac functions. Indeed, ECG quantifiers (e.g., heart rate variability or RR intervals) are suggested to be more strongly pertinent to the underlying pathological mechanisms in previous investigations. Then, ITD is imposed upon RR intervals to derive the proper rotations and unravel instantaneous TFE characteristics. Finally, intrinsic energy, intrinsic frequency entropy and intrinsic time entropy are extracted from the TFE characteristics for AF identification.

Keywords: Atrial fibrillation, intrinsic time-scale decomposition, intrinsic entropy

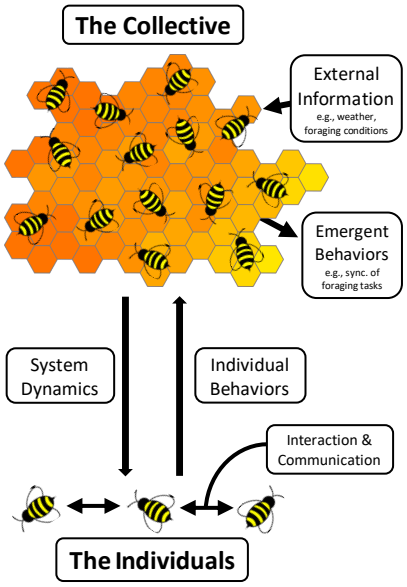
[1] Corresponding author: ccheng@binghamton.edu

What the Buzz Is All About: *Apis mellifera* as a Model for Collective Intelligence

Megan R. Chiovaro <megan.chiovaro@uconn.edu>
Alexandra Paxton <alexandra.paxton@uconn.edu>
Department of Psychological Sciences, University of Connecticut
Center for the Ecological Study of Perception and Action, University of Connecticut

Social insects provide a remarkable avenue for studying collective intelligence. Often referred to as "superorganisms," these multi-agent biological systems coordinate activity through interactions among the individuals, the collective, and the physical surroundings (see figure). One vital social invertebrate—*Apis mellifera*, the Western honeybee—exhibits multiscale individual and collective dynamics, and has long been used for the study of collective intelligence.

We propose a new approach to the study of collective intelligence in honeybees: an *ecological* approach, dedicated to the discovery of lawful animal-environment relationships [1] that can broaden our knowledge of collective dynamics and inform our approach to studying human behavior. Investigating underlying behavioral rules governing the colony will help us not only understand how the lack of a central controller can lead to more productive processes but will also lay the ground for creating effective human collectives. To that end, we outline potential paths for studying the collective behavior of honeybees and humans and present our current work toward bridging these domains.



One of the honeybees' most interesting adaptive systems is task allocation. Using vibrational and olfactory signals, forager bees encourage in-hive task productivity—such as food storage—to synchronize with increased profitability of foraging conditions [2]. Using a large-scale interactive online experiment platform [3], we will be simulating a task chain modeled after the system used by honeybees. Specifically, we will allow multiple participants to interact in real-time, creating an actual human collective, to work together to achieve a common goal. We hypothesize that, under the right conditions, humans will naturally perform task-switching in order to synchronize productivity of various roles, just as honeybees do.

Another honeybee phenomenon occurs when over-packed hives split, leading half of the colony to relocate. Scout bees are sent out in search of suitable locations and return to "advertise" candidates through vigorous dances, beginning a drawn-out competition between suitable sites. This phenomenon sparked the creation of dynamical systems models, simulating inter-organism communication such as neurons in bilateral decisions [4]. We are working toward an extension of this model, incorporating late-onset third choices with varying profitability. Doing so will help identify when human groups will ignore superior alternatives in group decision-making, solely because they were introduced "too late" in the process.

Further empirical investigation and modeling could improve our understanding of complex collective phenomena, both in social insects and humans. We encourage researchers to join us in taking a dynamical systems and ecological approach to understanding the compelling social behaviors of honeybees to inform the study of collective intelligence in ecology and psychology.

References

[1] Gibson (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
[2]Schneider & Lewis (2004). The vibration signal, modulatory communication and the organization of labor in honey bees, Apis mellifera. *Apidologie, 35*, 117-131.
[3] Paxton, Morgan, Suchow, & Griffiths (2018). Interpersonal coordination of perception and memory in real-time online social experiments. *Proceedings of the Cognitive Science Society*.
[4] Reina, Bose, Trianni, & Marshall (2018). Psychophysical laws and the superorganism. *Scientific Reports, 8*(1).

Exploratory Study of the Application of Anomaly Detection Methodology to the Analysis of Fatigue Data

Jakob Zeitler,[1] Chilukuri Mohan[1], Volker Weiss[2], Robert Haynes[2]

[1] Syracuse University
[2] Army Research Laboratory

Extended Abstract: A machine learning evaluation of a dataset resulting from Three-Point bending fatigue tests on glass-epoxy composites is presented. The test results are contained in a publication by Robert Haynes et al. [1] of the US Army Research Laboratory. Displacement versus number of cycles were recorded for four load ranges on the test-machine for 222 specimens. A small number of these were subjected to analysis that aims to support the theoretical understanding of the fatigue process. The results show that the compliance (see Figure 1), displacement over cycling load, increases to failure in three phases with a minimum compliance increase rate over the major portion of the life. This increase consists of at least two characteristics, a steady state monotonous increase interrupted at random intervals by short periods of significantly higher displacement increase steps. Classification algorithms, i.e. neural networks, support vector machines and random forests, show the prediction feasibility of failure via features from data of the last 50 cycles available. The lower displacement and the range of displacement are identified as most informative, leading to a classification AUC of 98%. Oversampling is used to reduce the rate of false negatives. More work, including the use of sensor fusion with other damage indicators, (e.g. acoustic emission) is suggested.

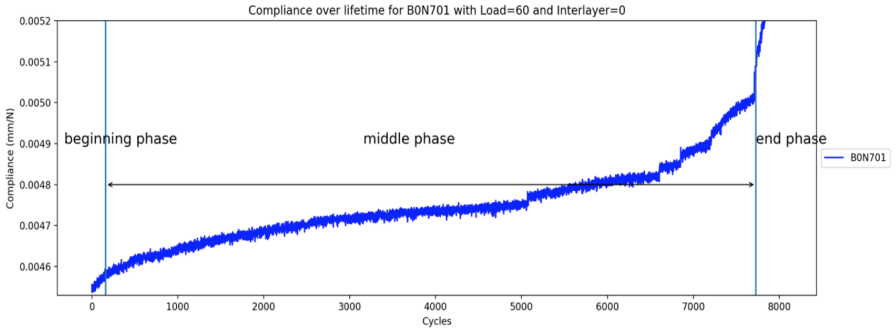


Figure 1: A compliance curve

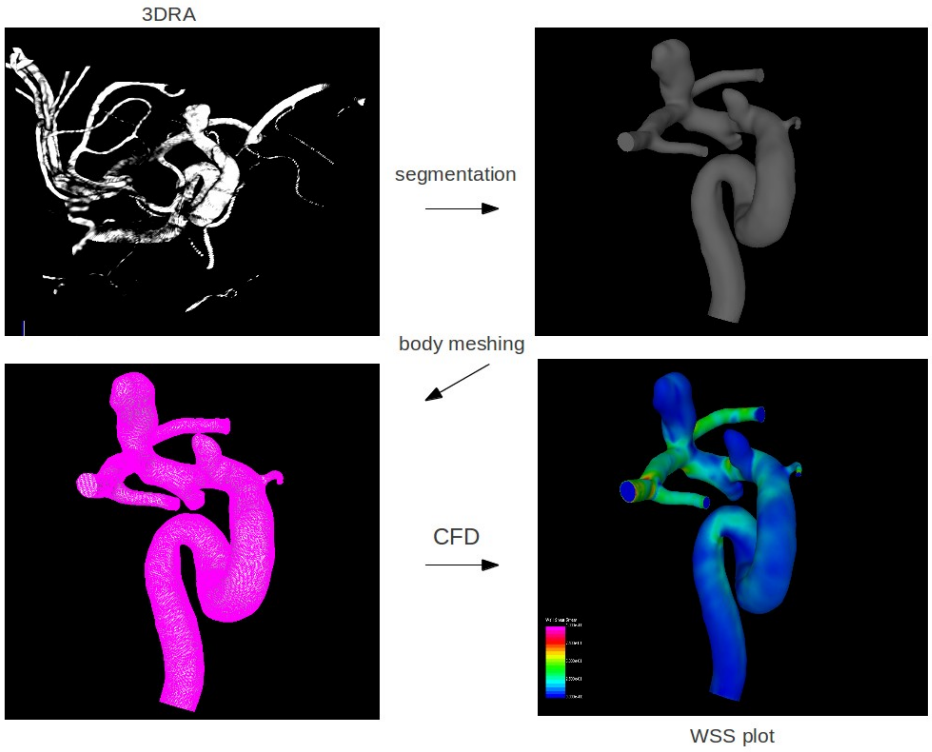
References

[1] Haynes, R; Boyd, S; Lawrence, B: Compression-after-Impact and Bend Fatigue Results of Glass/Epoxy Composites with Compliant Interlayer and Needling Interlaminar Enhancements. US Army Research Laboratory Aberdeen Proving Ground United States (2016)

CFD Study to Identify Risk Factors of Cerebral Aneurysm Rupture

It is widely known that the rupture of intracranial aneurysms has devastating consequences with large morbidity and mortality rates. The etiology of aneurysms has been discussed earlier by many scientists and clinicians. It is now well accepted that hemodynamic environments in aneurysms play a key role on initiation, growth, and rupture processes in combination with biochemical and physical interactions. Computational fluid dynamics (CFD) using patient specific aneurysms from 3DRA or CTA images has been used for several years to identify mechanical risk factors (wall shear stress (WSS), oscillatory shear index (OSI), flow structure, size, or aspect ratio (AR), etc) associated with aneurysm evolution and rupture. It has been also successfully employed to understand flow characteristics before and after surgical treatments using minimally-invasive endovascular devices such as coils and flow diverters (FD) in order to help the clinical decision making process. However, we find that due to i) the large number of CFD pre-processing inputs such as wall material properties, blood viscosity, mesh resolution, boundary conditions and geometry and ii) the lack of knowledge of the detailed vascular remodeling processes, further efforts of validation and reproducibility of CFD with larger clinical datasets and of fluid structure interaction (FSI) analyses are still essential to identify standardized mechanical risk factors. In this talk, I present a FSI study to investigate the relationship between the intra-aneurysmal hemodynamic conditions and wall mechanical properties in eight unruptured aneurysms, resected and mechanically tested for their failure using patient specific CFD models.

Keywords: Computational Fluid Dynamics, Intracranial Aneurysm, Nonlinear Dynamics



Leader Emergence and Collective Action during Crisis

Anwesha Choudhury, Ankita Kulkarni, Shalini Kurumathur, Savisesh Malampallayil, Sriniwas Pandey
State University of New York, Binghamton

Abstract

On September 11, 2001, the passengers of Flight 93 displayed courageous collective action when they counter hijacked the flight from terrorists. In this paper, we attempt to simulate the events that are known to have taken place on the day, with the aim to understand emergent properties of the system. We utilize Agent Based Modeling and build a two-stage model. In the first stage, leader emergence takes place through localized agent interactions based on presence of different characteristics of the agents: extraversion, prosocial behavior, and task expertise. In the second stage, we model two processes: globalized agent interaction and impact of external information. The agents have two dispositions: risk avoidant and action oriented. We operationalize collective action as there being a majority of action-oriented agents. Using multiple simulations, our results show that information and prosocial behavior play fundamental roles for emergence of collective action. The assumptions, implications and future research are discussed.

Keywords: leader emergence, collective action, task expertise, extraversion, prosocial behavior

1. Introduction

Imagine that you are aboard a flight and while the flight is airborne you get to know that the flight has been hijacked by terrorists. A plethora of thoughts and emotions would start cruising through your mind. What would your first reaction be - would you panic and leave things to fate, or would you exhibit courage and start planning for a resolution? Would you demonstrate leadership and guide others out of this problem or would you look for support and follow someone who tries to solve this problem? Would you try solving the crisis all by yourself or would you encourage others to approach the problem collectively? On September 11, 2001, passengers on Flight 93 were puzzled with similar questions in their minds when they learnt of their flight hijack [1]. When confronted with the urgency of their situation, the passengers and crew chose to act heroically and demonstrated a collective behavior to organize a counterattack on the hijackers.

We deal with many crisis situations in our day-to-day life, be it individual-level situations such as job loss, diseases and family disputes, or community-level situations such as natural disasters, stock market crashes and crime related incidents. From an organizational perspective, crisis is an unusual event, which may have unknown cause and effect, with high ambiguity and low probability of occurrence [2]. When confronted with crisis, people would either resign to their fate and accept failures, or step up to resolve the crisis through rapid responses. Crisis resolution is very different from solving other types of problems since people lack an understanding of the cause and consequences of these events as a result of which unpredictability in their behavior increases [3]. Hence research on crisis and its resolution has drawn a lot of attention.

2. Theoretical Background

Of particular interest to us is the phenomenon of leader emergence and collective action during crisis. Our model draws inspiration from the three narratives proposed by Quinn and

Worline [1] who use content analysis to study the courageous collective action taken by the people onboard the flight 93 on 9/11- a personal narrative that helps them understand who they are beyond the immediate situation and manage the intense emotions that accompany duress, a narrative that explains the duress that has been imposed upon them sufficiently to make moral and practical judgments about how to act, and a narrative of collective action—and the resources that make the creation of these narratives feasible

Extensive research has been carried out related to many attributes of leader emergence such as personality traits [4][5][6]cognitive traits [7][8][9]and traits related to emotions [10][11]. However, leadership emergence during crisis and collective action to solve the crisis has been under-researched. We draw on the story of the courageous collective action from Flight 93 and develop a model to understand the leader emergence in crisis using three main antecedents of leader emergence: extraversion, prosocial behavior and task expertise.

Extraversion*:* The five-factor model of personality has been extensively used to study leader emergence. Extraversion, which represents the tendency to be sociable, assertive, active, and to experience positive effects, such as energy and zeal, has been observed to play an important role in leader emergence [12]. Those individuals who have a high level of energy and stamina, and are generally active are more likely to emerge as leaders [13][14]. We consider this personality trait in our model since crisis situation needs action-oriented people who are active to the situation and assertive in terms of convincing others and implementing strategies.

Prosocial Behavior: Prosocial behavior has been defined as any behavior consisting of helping, shame, cooperating and volunteering that is carried out to produce and maintain the well-being and integrity of others [15]. Leaders who exhibit prosocial behavior often engage in personally risky behavior and motivate follower prosocial behavior which benefits the collective [16]. One of the most consistent observations reported by pioneer field researchers was that during the crisis period of disasters, a prosocial behavior helped immensely in coping with the extreme and unusual demands of a disaster situation [17].Hence, we have considered this to be an important variable in our model.

Leadership Emergence: Leadership emergence is a result of the consensual agreement in the perceptions of fellow group members that one individual is the leader in a group with no appointed leaders [18]. Emergent leaders of such teams exhibit significant influence over other group members even though no formal authority has been vested on them [19]. Emergence of leaders also is a critical factor in resolving crisis situations [20].

Task Expertise: Various studies from literature have revealed that leaders who have greater cognitive abilities in terms of intelligence, planning skills, task expertise, and creative thinking skills are able to make better decisions by resolving complexity and ambiguity [21][22][23][24][25][26]. Task expertise stems from cognition and has been defined as something personally encountered; knowledge, skill or practice derived from direct observation or participation in event" [27]. We consider task expertise as an important variable in our model since leadership emergence requires relevant task experience [28]. It is also an important factor in resolving crisis [29].

3. Method

Agent Based Simulation was used as we needed to model two hundred discrete individuals who had complex properties and behavioral rules [30].

Rules

Leader Emergence. First stage of the model- Following are the possible combinations for leader emergence:

Task Expertise	Extraversion	Prosocial behavior
1	1	0
1	0	1
0	1	1
1	1	1

Collective Action. Second stage of the model- It is important to note that this stage occurs only after leaders have emerged. The second stage is divided into two following phases:

Local neighbors' interaction- When a group of 2-5 neighbors interact, only the presence of a leader in the group will change the disposition of the agents from risk avoidance to action-orientation.

Global neighbors' interaction- Once the small groups are transformed, they interact with larger groups of 12-17 agents. Based on the general interaction, the following rule was created:

- Presence of two leaders in the group of 12-17 will change the disposition of all the agents from risk avoidant to action-orientation

During this interaction, crisis related information from external source (for e.g. in the flight 93, the passengers were able to connect to their families using the Wi-Fi on their phones, through this connection, they were able to garner information regarding the initial crisis) plays an important role, depending on the level of information (low/high), the following rules are created:

When information is high:

- If a 12-17 group constitutes of 35% of agent population who are action oriented, then the entire group turns into action-orientation individuals.
- If there is a leader in the 12-17 agent group, 45% of the agent population turns into action-orientation individuals

When information is low:

- If a 12-17 group constitutes 60% of agents who are action-oriented, then the entire agent population will transform into action-oriented individuals

If there is a leader in a group, 18-20% of agent population will turn into action-oriented agents

Assumptions

The following are some of the assumptions of our model:

1) In our study we use agent's extraversion, prosocial behavior and task expertise as important antecedents for leader emergence. There may be many other cognitive, emotional and personality related characteristics which may play important roles in crisis leader emergence, but we focus on the above three antecedents to understand a larger picture, but we focus on the above three antecedents to understand a larger picture. Extraversion has been associated with leader emergence consistently. The average correlation of extraversion with leader emergence is 0.3 [36][38]. In the current situation where a simulation of flight hijack is portrayed, the passengers (aka agents) who have prosocial behavior are prone to offer comfort and help to others [33] and hence these people emerge as the support pillars or leaders for the collective. Empirically the correlation of prosocial behavior with leader emergence is 0.10 [39]. Task expertise or having relevant prior knowledge regarding the task at hand is another essential element to be considered in

managing a crisis [35][37]. Professionals such as army men, doctors, or firemen may have prior experience dealing with such crisis and can maintain calm and encourage others during these moments, hence we consider task expertise as a leader predictor variable. The empirical correlation of task expertise with leader emergence is 0.27 [35].

2) We use a combination of two out of three characteristics (task expertise, prosocial behavior and extraversion) as input conditions for leader emergence which means that if an agent is only extroverted then he/she will not emerge as leader whereas if an agent is extroverted and prosocial then he/she will emerge as leader. All three characteristics are equally weighted.
3) The proportion of agents having relevant task expertise is low (2%), agents having extraversion is moderate (50%) whereas prosocial behavior of agents is manipulated. Research has shown that extraversion is a stable trait and is present in a majority of the population, and task expertise is based on previously accumulated knowledge due to which only very few passengers may be well equipped to deal with crisis situations. Thus, only prosocial behavior is manipulated in our model.
4) Crisis relevant information is considered to be essential for collective action because agents who have sufficient amount of relevant knowledge regarding the crisis at hand will be able to transform themselves more readily into action-oriented agents as they may be more confident that they will be able to produce a good result [32].
5) Each agent is an independent entity which means he/she travels alone on the flight. Although this is artificial, it was hard to replicate groups of agents and their interactions and the impact of the interactions on leader emergence (for e.g. if a family traveled together and if the father realized that his daughter may be in danger then irrespective of his behavioral traits, he may still emerge as a leader), hence we did not consider group of agents. Previous papers on ABM methodology have said that the goal of the method is to define classes of agents rather than describing all possible behavioral aspects [31].
6) We defined action-orientation as a willingness to counter-hijack the plane. While there may have been other actions possible under the circumstance, we chose this since we were following a known historical occurrence.
7) We assumed that leaders or presence of action-oriented individuals influence others to change their behavior from risk avoidant to action-orientation. Action-orientation is defined as an attribute associated to individuals of, they concentrate on a specific goal and take risks [43]. In their research they demonstrated that an action-oriented basketball player takes more risks to shoot faster and more often to the basket. Studies in organizational set ups have been conducted to study the association between action-orientation and risk-taking behavior with similar outcomes [44][45][46]. Since leaders can influence individuals becoming action-oriented, we apply the simple majority rule [47] in our study to assume that collective action will emerge if at least 50% of agents will demonstrate this switching behavior from risk-avoidance to risk-taking.
8) Quinn & Worline reported a vote being taken before the passengers started the counter-hijacking [1]. Based on this, we have assumed in our model that 50% of agent population changing from risk avoidant behavior to action-orientation would lead to collective action being undertaken.
9) All leaders are assumed to be action-oriented, this has support from literature as leaders are the ones who partake in decision making and are central to the dynamic changes that occur in a situation. In the face of a problem, leaders must pay attention to their own feelings as

well as the feelings of those around, develop a vision, and turn them into action to demonstrate that they are serious about their vision [48]. Previous research has shown that people like their leaders to be strong and action-oriented, and for any change to be effective, leaders need to create visions that guide people's actions and convert the same into teamwork [49][50]. When face with a crisis, we argue that resolving the same is a worthwhile goal and leaders would develop a vision to resolve the situation. Also, leaders will influence their followers to demonstrate the action-oriented competence to perform tasks and activities necessary to resolve the crisis.

10) Quinn &Worline reported in their paper that many people were surprised by the radical notion proposed by a passenger regarding hijacking being a suicide mission [1]. Based on this, we assume that very little information is present in the local neighbors' group (2-5 agents) and hence this small amount of information does not play a role in changing the disposition of agents from risk avoidant to action oriented, so here information is considered to be insignificant.
11) Initially we assumed that 10% of the agent population was action-oriented whereas 90% of the agent population was risk avoidant. This is in keeping with collective action literature which presumes that initially fewer members of a population are disposed towards taking action [42].
12) To simplify the model, demographic characteristics such as gender, age, race are not considered.

4. Simulation and Experimental results

To simulate the model, "PyCX simulator" simulator is used (Figure 1). It is a python-based simulator for complex systems. The simulator can show real-time states and state transitions of the system. We are running our experiment to show output in two different perspectives; first is to mimic the flight 93 situation and another is to study the effect of different human characteristics on the leader emergence and collective decision-making process. As very limited parameters are in our control, we are mainly showing how manipulating the distribution of people with high prosocial behavior affects the outcome. Another insight, we are interested in, the role of some external information in the crisis situation.

Figure 2 depicts the experimental outcomes. As the number of individuals with high prosocial behavior (psb) in the crowd increases, the probability of emerging a leader increases and that leads to the conversion of a large proportion of people to action-oriented mode. The presence of some external information increases the conversion, i.e. even for the smaller number of high prosocial people, we get more action-oriented people.

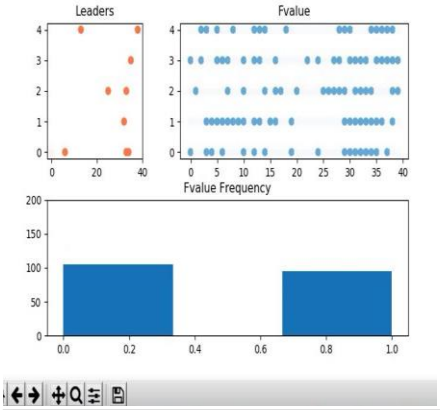


Figure 1. PyCX simulator

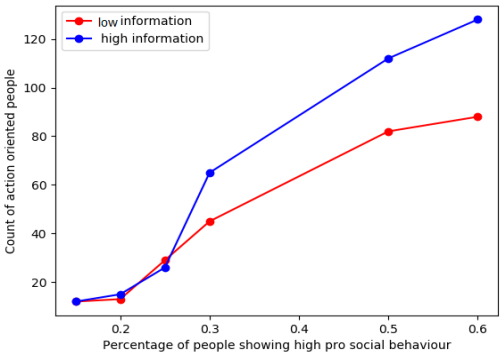


Figure 2. Effect of information and pro social behavior

5. Discussion

An important aspect of the model is the relevance placed on agents' disposition of being risk avoidant versus action-oriented. Based on our assumptions, first, our simulation shows that presence of two characteristics out of three: task expertise, extraversion and prosocial behavior will lead to leader emergence. Once these leaders emerge, they convert the risk avoidant agents into action-oriented in small groups, at a later stage the level of information present (low/high), leaders and the percentage of risk avoidant versus action-orientation aids in emergence of collective action. Although we made numerous assumptions, we had an important finding: prosocial behavior or the ability to be concerned for the social well-being of others is more important than extraversion in a crisis situation. This follows logically from the notion that people who are more inclined to help others will take action in a risky situation. This was achieved through manipulation of the proportion or number of agents displaying prosocial behavior in the flight. Although our model was loosely based on flight 93 crisis, we were able to reproduce leader emergence and collective action as represented in this crisis situation.

6. References

[1] Quinn, R. W., & Worline, M. C. (2008). Enabling courageous collective action: Conversations from United Airlines flight 93. *Organization Science*, *19*(4), 497-516.

[2] Pearson, C. M., & Clair, J. A. (1998). Reframing crisis management. *Academy of Management Review*.

[3] Mumford, A. (2006). Management Development-strategies for action. *Development and Learning in Organizations: An International Journal*, 20(2).

[4] Kirkpatick, S. A., & Locke, E. A. (1991). Leadership: do traits matter?. *Academy of Management Perspectives*, 5(2), 48-60.

[5] House, R. J., & Aditya, R. N. (1997). The social scientific study of leadership: Quo vadis? *Journal of Management*, 23, 409–473.

[6] Shamir, B., & Howell, J. M. (1999). Organizational and contextual influences on the emergence and effectiveness of charismatic leadership. *The Leadership Quarterly*, 10(2), 257-283.

[7] Kirkpatick, S. A., & Locke, E. A. (1991). Leadership: do traits matter?. *Academy of Management Perspectives*, 5(2), 48-60.

[8] Daft, R. L. (1999). Leadership: Theory and practice. Orlando, FL: *Dryden Press*.

[9] Ford, C. M., &Gioia, D. A. (2000). Factors influencing creativity in the domain of managerial decision making. *Journal of Management*, 26(4), 705-732.

[10] Kellett, J. B., Humphrey, R. H., & Sleeth, R. G. (2006). Empathy and the emergence of task and relations leaders. Leadership Quarterly.

[11] Pescosolido, A. T. (2002). Emergent leaders as managers of group emotion. *Leadership Quarterly*.

[12] Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: a qualitative and quantitative review. *Journal of applied psychology*, *87*(4), 765.

[13] Kirkpatick, S. A., & Locke, E. A. (1991). Leadership: do traits matter?. *Academy of Management Perspectives*, 5(2), 48-60.

[14] Ensari, N., Riggio, R. E., Christian, J., & Carslaw, G. (2011). Who emerges as a leader? Meta-analyses of individual differences as predictors of leadership emergence. *Personality and Individual Differences*, *51*(4), 532-536.

[15] McNeely, B. L., & Meglino, B. M. (1994). The role of dispositional and situational antecedents in prosocial organizational behavior: An examination of the intended beneficiaries of prosocial behavior. *Journal of applied psychology*, *79*(6), 836.
[16] De Cremer, D., Mayer, D. M., Van Dijke, M., Schouten, B. C., & Bardes, M. (2009). When does self-sacrificial leadership motivate prosocial behavior? It depends on followers' prevention focus. *Journal of Applied Psychology*, *94*(4), 887.
[17] O'Reilly, C. A., & Chatman, J. (1986). Organizational commitment and psychological attachment: The effects of compliance, identification, and internalization on prosocial behavior. *Journal of applied psychology*, *71*(3), 492.
[18] Berdahl, J. L. (1996). Gender and leadership in work groups: Six alternative models. *Leadership Quarterly*, *7*(1), 21–40.
[19] Schneider, C. E., & Goktepe, J. R. (1983). Issues in emergent leadership: The contingency model of leadership, leader sex, leader behavior. In H. H. Blumberg, A. P. Hare, V. Kent, & M. F. Daview (Eds.), *Small groups and social interactions* (Vol. 1).
[20] Pillai, R. (1996). Crisis and the Emergence of Charismatic Leadership in Groups: An Experimental Investigation 1. *Journal of Applied Social Psychology*, *26*(6), 543-562.
[21] Cornwell, J. M. (1983). Meta-analysis of selected trait research in the leadership literature. In *meeting of the Southeastern Psychological Association, Atlanta, GA*.
[22] Lord, R. G., De Vader, C. L., & Alliger, G. M. (1986). A Meta-Analysis of the Relation Between Personality Traits and Leadership Perceptions. An Application of Validity Generalization Procedures. *Journal of Applied Psychology*.
[23] Bass, B. M. (1990). Bass and Stogdill's handbook of leadership. *New York: Free Press*.
[24] Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: a quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, *89*(3), 542.
[25] Marta, S., Leritz, L. E., & Mumford, M. D. (2005). Leadership skills and the group performance: Situational demands, behavioral requirements, and planning. *The Leadership Quarterly*, *16*(1), 97-120.
[26] Mumford, M. D., Friedrich, T. L., Caughron, J. J., & Byrne, C. L. (2007). Leader cognition in real-world settings: How do leaders think about crises?*The Leadership Quarterly*, *18*(6), 515-543.
[27] Jacoby, J., Troutman, T., Kuss, A., & Mazursky, D. (1986). Experience and expertise in complex decision making. *ACR North American Advances*.
[28] Mumford, M. D., Scott, G. M., Gaddis, B., & Strange, J. M. (2002). Leading creative people: Orchestrating expertise and relationships. *The leadership quarterly*, *13*(6), 705-750.
[29] Farazmand, A. (2007). Learning from the Katrina crisis: A global and international perspective with implications for future crisis management. *Public Administration Review*, *67*, 149-159.
[30] Sayama, H. (2015). *Introduction to the modeling and analysis of complex systems*. Open SUNY Textbooks.
[31] Rand, W., & Rust, R. T. (2011). Agent-based modeling in marketing: Guidelines for rigor. *International Journal of Research in Marketing*, 28(3), 181-193.
[32] Liu, B. F., Austin, L., &Jin, Y. (2011). How publics respond to crisis communication strategies: The interplay of information form and source. *Public Relations Review*, 37(4), 345-353.

[33] Batson, C. D., & Powell, A. A. (2003). Altruism and prosocial behavior. Handbook of psychology, 463-484.
[34] Costa Jr, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. Personality and individual differences, 13(6), 653-665.
[35] De Souza, G., & Klein, H. J. (1995). Emergent leadership in the group goal-setting process. Small group research, 26(4), 475-496.
[36] Ilies, R., Gerhardt, M. W., & Le, H. (2004). Individual differences in leadership emergence: Integrating meta-analytic findings and behavioral genetics estimates. International Journal of Selection and Assessment, 12(3), 207-219.
[37] Hawkins, K. W. (1992). Effects of Gender and Communication Content on Leadership Emergence in Small Task-Oriented Groups. Small Group Research, 26(2), 234-249
[38] Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: a qualitative and quantitative review. Journal of applied psychology, 87(4), 765.
[39] Mitchell, T. D., &Bommer, W. H. (2018). The interactive effects of motives and task coordination on leadership emergence. Group Dynamics: Theory, Research, and Practice, 22(4), 223.
[40] Taggar, S., Hackew, R., &Saha, S. (1999). Leadership emergence in autonomous work teams: Antecedents and outcomes. Personnel Psychology, 52(4), 899-926.
[41] Smith, J. A., &Foti, R. J. (1998). A pattern approach to the study of leader emergence. The Leadership Quarterly, 9(2), 147-160.
[42] Ostrom, E. (2000). Collective action and the evolution of social norms. Journal of economic perspectives, 14(3), 137-158.
[43] Raab, M., & Johnson, J. G. (2004). Individual differences of action orientation for risk taking in sports. Research quarterly for exercise and sport, 75(3), 326-336.
[44] Kuhl, J. (1984). Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action control. In Progress in experimental personality research (Vol. 13, pp. 99-171). Elsevier.
[45] Singh, J. V. (1986). Performance, slack, and risk taking in organizational decision making. Academy of management Journal, 29(3), 562-585.
[46] Anderson, C., &Galinsky, A. D. (2006). Power, optimism, and risk-taking. European journal of social psychology, 36(4), 511-536.
[47] Inada, K. I. (1964). A note on the simple majority decision rule. Econometrica: Journal of the Econometric Society, 525-531.
[48] Goleman, D., Boyatzis, R. E., & McKee, A. (2002). The new leaders: Transforming the art of leadership into the science of results.
[49] Sandmann, L. R., & Vandenberg, L. (1995). A framework for 21st century leadership. Journal of Extension, 33(6), 1-9.
[50] Tichy, N. M., & Cohen, E. (1997). The leadership engine. New York, NY: HarperCollins.

Interpreting biomarkers of milk immunity

ML Duris[1], M Manganello[1], C Torres[1], ME Gauck[1], S Li[1], D Armstrong[1], T Hopt[1], R Spathis[1], K Wander[1] [1]Department of Anthropology, Binghamton University (SUNY), Binghamton, NY

How the immune system of milk protects infants against infectious disease and guides their immune development is an important focus recent research. This research is complicated, however, because milk immune content is much easier to measure than milk immune activity. Milk immune activity results from interactions between many immune factors in milk and may not be predictable from levels of any one factor. We attempted to validate two biomarkers of milk immunity, one pro- and one anti-inflammatory, against *in vitro* milk immune responses.

Volunteers provided milk specimens, which were incubated in mammalian cell-culture medium with a bacterial stimulus (*Salmonella*, a pathogen, or *Bifidobacteria*, a commensal organism) at 37C in an anaerobic environment for 24 hours, and baseline and stimulated specimens were evaluated via enzyme immunoassay. Outcomes of interest ("responses") were increases from baseline in two pro-inflammatory cytokines, interleukin (IL)-6, and interferon (IFN)-γ and one anti-inflammatory cytokine, IL-10. Predictors of interest were two biomarkers, IL-6 (pro-inflammatory) and IL-1 receptor antagonist (IL-1ra, anti-inflammatory). Use of IL-1ra, in particular, is increasingly as a biomarker of the anti-inflammatory capacity of milk.

Paired baseline and stimulated responses were available for 25 specimens for IL-1ra (geo. mean IL-1ra, 243.1 pg/ml; range: 25.1, 1318.3), and 31 specimens for IL-6 (geo. mean IL-6: 4.2 pg/ml; range: 3.3, 24.1). Three types of response were apparent: none (19%), an increase in IL-6 (71%), and an increase in all three cytokines (10%). These differed by baseline IL-6, but not IL1-ra (see Figure 1). The magnitude of IL-6 increase was unassociated with baseline IL-6 or IL-1ra.

Identifying anti-inflammatory responses was not straightforward: when IL-10 increased from baseline, it was often in concert with increases in other cytokines. The anti-inflammatory biomarker, IL-1ra, was *positively* associated with IFN-γ responses and IL-10 responses. However, for responses to *Salmonella*, IL-1ra was also positively associated with the ratio of IL-10 to IL-6: for each 1 pg/ml increase in IL-1ra, the ratio shifted toward IL-10 (p=0.011) by 0.2%.

Overall, we found limited support for use of IL-6 or IL-1ra as biomarkers of milk immunity: Higher milk IL-6 was associated with more robust immune responses. Baseline milk IL-1ra could not distinguish those milk specimens that exhibited no response to stimuli, but was associated, when a response occurred, with proportionately more anti-inflammatory activity. Given the complexity of immune regulation, and increasing interest in the immune system of milk, careful validation of biomarkers of milk immunity is clearly warranted.

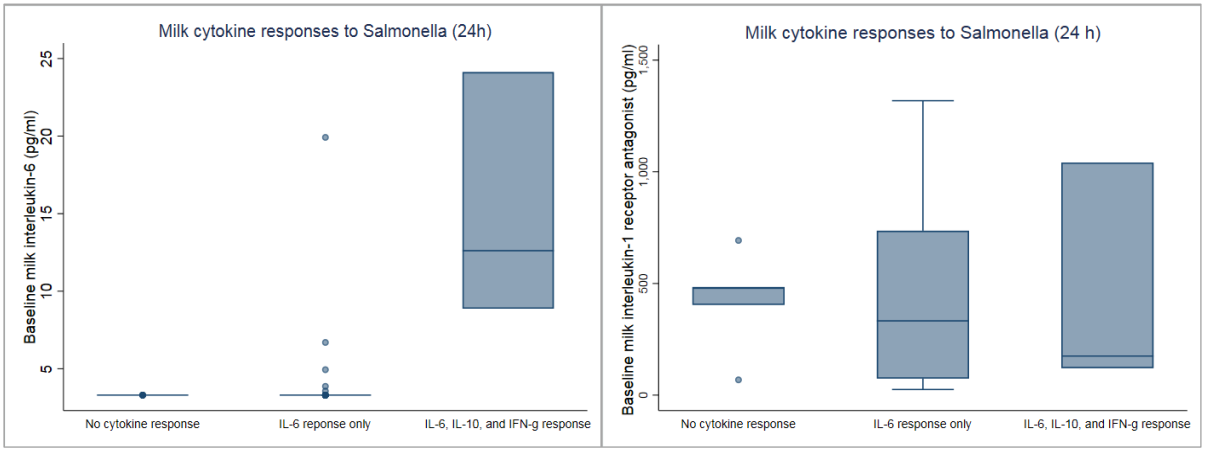


FIGURE 1. Milk cytokine responses to Salmonella over 24hrs associated with baseline IL-6 (left) and IL-1ra (right) levels.

INFERRING THE "SOCIAL NETWORK" OF CELLS FROM ESTIMATION OF INFORMATION TRANSFER AND GRAPH THEORY

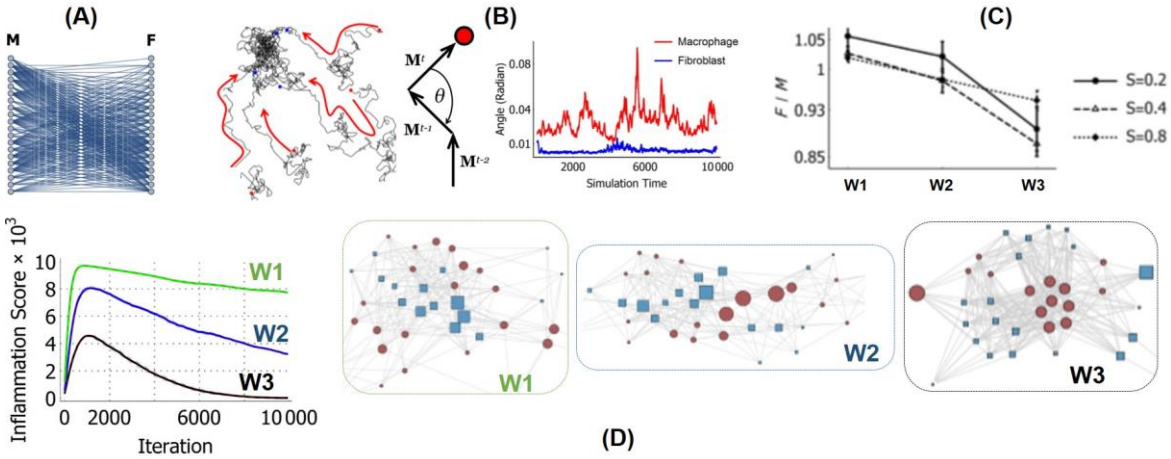
Israr B M Ibrahim[1], Ramana M Pidaparti[1]
[1]College of Engineering, University of Georgia, Athens, GA, 30602

Coordinated movement of cells play important role in inflammation and wound healing. In collective cell migration, some cells dominate the coordinated movements of the collective (the leader cells), determining the direction and velocity of the rests (the followers) [1]. In this study, we propose that the time course of wound healing is dependent to the relationship between the dominating cells and the follower cells. This relationship can be modeled as a network, akin to a social network. This model enable one to compute the centrality of the network, thus establishing which cells dominate the collective migration of cells. We used simulated data of wound healing with different time course (Fig 1(D)) that are progressively worst (in terms of recovery and wound healed at the end of the time course). The simulated data were generated from an agent-based model that models the migration and interaction of macrophages and fibroblasts during wound healing. To determine the network of cell relationship, we considered a network between macrophage and fibroblast (Fig 1(A)). To determine the connection between a macrophage (M) and fibroblast (F), we computed transfer entropy (TE) between a pair of M and F. The M and F migration aspects considered for TE was the change of angle with respect to time, $\theta(t)$ (Fig 1(B)). Next, the M-F networks were generated according to:

$$f(M_i, F_i) = \begin{cases} \{M_i, F_i\} \in G, & TE_i > T_c \\ \{M_i, F_i\} \notin G, & TE_i \leq T_c \end{cases}$$

where G is the graph representing M-F network (Fig 1(A)), $\{M_i, F_i\}$ denotes an edge between a macrophage node M_i and fibroblast node F_i, TE_i is the transfer entropy of a pair i, and T_c is a threshold. The resulting M-F networks were computed for their betweenness centrality. Results indicate that the time course of wound healing can be distinguished by their corresponding M-F networks. Partial wound healing is dominated by macrophages (with M nodes having higher centralities), and effective wound healing process is dominated by fibroblasts (the MF networks prefer higher centralities on the F nodes).

Keywords: Graph Theory, Collective Dynamics, Wound Healing



(A) A model of relationship between macrophage (M) and fibroblast (F) (MF network). (B) To build the MF network, transfer entropy between pairs of M and F was calculated based on their change of angle over time. (C) The computed centralities show decreasing dominance of M over F as

wound healing worsened (the effectiveness of wound healing is W1> W2 > W3). (D) The three wound healing time courses (left) and samples of corresponding MF networks that represent the collective dynamics of macrophages and fibroblasts.

[1] Haeger, Anna et al., Collective cell migration: guidance principles and hierarchies, Trends in Cell Biology , Volume 25 , Issue 9 , 556 – 566, doi: 10.1016/j.tcb.2015.06.003

Network Characteristics of Bronchoconstriction in Lung Airways: A Computational Study

Israr B M Ibrahim[1], Ramana M Pidaparti[1]
[1]College of Engineering, University of Georgia, Athens, GA, 30602

Bronchoconstriction (such as in the case of asthma) is a complex process of inflammatory response resulting in inhomogeneous (patchy) distribution of lung ventilation. In this study, we demonstrate that the bronchoconstriction process is analogous to network optimization and adaptation [1]. The lung is a naturally-occuring network that consists of airway units connected in hierarchical fashion, forming a tree network (Fig 1(A)). As a consequence, we may use the network-based model to quantify the network properties of diseased airways network and quantify the effect of different disease to airway morphology. The model for the airway network adaptation is a coupled equation,

Flow: $\mathbf{Lp} = \mathbf{q}$

Adaptation: $\frac{d\mathbf{K}}{dt} = \mathbf{q}^{2\gamma} - \mathbf{K} + \kappa\, e^{\left(\frac{-t}{1+\rho}\right)}$

where $\mathbf{L} = \boldsymbol{A^T K A}$ is Laplacian matrix of the airway tree network which is calculated from $\mathbf{A}$ (the incidence matrix) and $\mathbf{K}$ (conductivity matrix), $\mathbf{p}$ is nodal pressure, $\mathbf{q}$ is nodal flow rate, γ is a constant, κ and ρ are control parameters. For low ρ and κ the system leads to dense connectivity, and high ρ and κ yields hierarchical networks. For bronchoconstriction, we found that setting $\rho = 0.625$ and $0.16 \leq \kappa \leq 3.2$ are sufficient to model bimodal ventilation of the airway networks (Fig 1(B), ventilation here is defined as $\mathbf{q}$ times airway area). Furthermore, the changes in network measures of bronchoconstricted airway networks (Fig 1(C)) were also computed. The airway network clustering (measured by community) and connectivity (measured by Wiener Index) shows declining trend as bronchoconstriction progressively worsened. This trend also holds when the airway network asymmetry level was increased by increasing variation (σ) of $\mathbf{K}$ across the airway's edges (Fig 1(C)). Furthermore, the Eigenvector centralities of diseased and non-diseased airway networks are clearly separable. Hence, the bronchoconstriction as a process may be analagous to network optmization/adaptation. In addition, transforming the airway morphology into network description may be a viable way to quantify the effect of bronchoconstriction and other diseases to the lung airways.

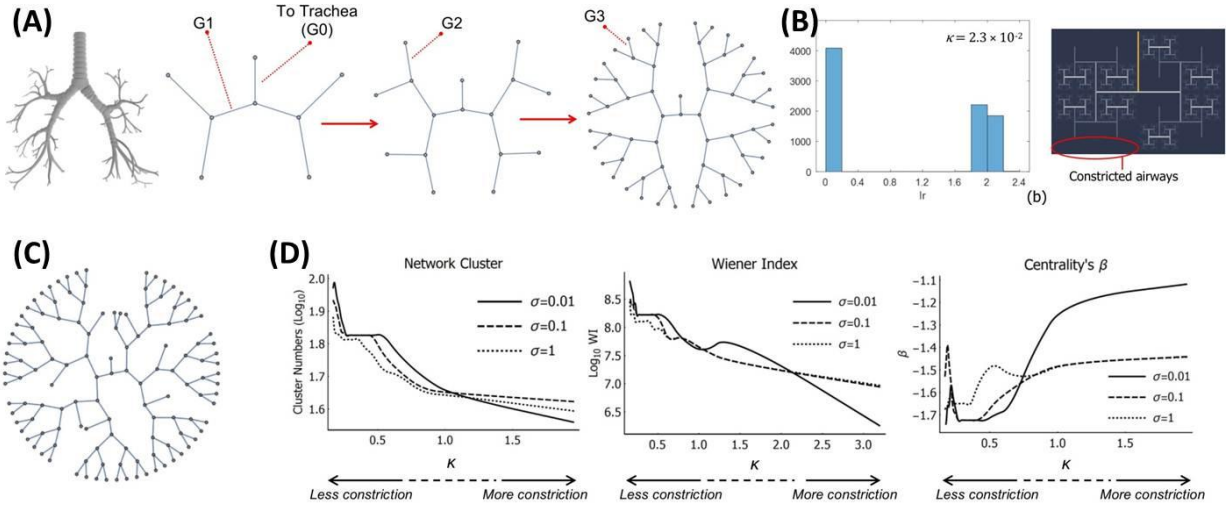


Fig 1. (A) The structure of the lung (left) can be represented as a network of airways (right). (B) Spatial distribution of ventilation in bronchoconstriction case is bimodal (left). (C) A sample of altered lung airways after bronchoconstriction. (D) The airway network under bronchoconstriction corresponds to reduced network complexity which leads to a more vulnerable airways network.

[1] H. Ronellenfitsch, E. Katifori, Global Optimization, Local Adaptation, and the Role of Growth in Distribution Networks, Phys. Rev. Lett. Vol 117, pp. 138301, Sept. 2016

The Evolution of Poverty Traps: An Agent-based Modeling Approach

Todd Guilfoos[1] and Emi Uchida[1]

Poverty traps are especially important in the developing world as the poorest of households struggle to make ends meet. Chronic poverty can exist for generations in some communities. One possible reason that people suffer from this chronic poverty is a lack of assets and capital that allow them to accumulate satisfactory living conditions. Often these at-risk populations also heavily rely on natural resources which can provide negative feedbacks to the poverty trap as they further deplete the natural resource.

We investigate the poverty dynamics of agents using common property resources through evolutionary game theoretic concepts to understand the sensitivity of poverty traps to shocks and selection. Evolutionary game theory has been used previously to understand cooperation in situations where the economic incentives seem to predict non-cooperative actions. We employ a similar framework to understand how thresholds and population structure affect selection of strategies when poverty traps exist. Are systems particularly sensitive to shocks common to all households (storms or contagious diseases) which lead to increased use of common property resources as social safety nets? Are systems sensitive to idiosyncratic shocks at the household level? Lastly, do initial conditions of population structure effect outcomes and how so. These aspects are thought to be important to poverty trap dynamics which can push and keep villagers beneath the poverty threshold.

Methods:

We use the structure of a game theoretic model which has a stock dependent payoff. This structure allows for the dynamics of extraction of the common property resource to effect future generations of the game. We discretize the strategy space to make computations more tractable in which households choose strategy $x \in \{H, L\}$ where high extraction is denoted, H, and low extraction is, L. Payoffs are contingent on group play of other agents in a gridded environment and local stock of the common pool resource. Therefore, the stage-game is similar to a prisoner's dilemma, yet the dynamics are subject to the population structure as well as the state of the natural resource.

In each generation agents reproduce based on fitness measure. Fitness is defined on expected payoffs of an agent and the intensity of selection, w. Mutation occurs with a probability of μ, where instead a random strategy is chosen for a randomly selected agent in the population.

This process is used to identify which strategies are chosen by agents and how poverty traps and shocks affect the sensitivity for so called 'good' equilibria vs. 'bad' equilibrium; put differently, how persistent is poverty to game structures. To the authors knowledge this is the first time that evolutionary game theory has been used to further understand poverty traps.

[1] University of Rhode Island, Environmental and Natural Resource Economics Department

Cities in the Rise and Decline of Civilizations

Cameron Harwick — The College at Brockport

Abstract. Why do societies rarely return to the technological frontier once they lose their place there? This paper uses a computational agent-based model to model urbanization as a time-inconsistent process endogenously driving both the rise and the decline of civilizations. Starting from a Malthusian rural economy, early-stage urbanization is driven by increasing returns to the local agglomeration of human capital. However, urbanization also raises the opportunity costs of the intergenerational transmission of that human capital. Economic growth continues so long as the network effects continue to draw human capital from rural areas, and the civilization stagnates once that pool is depleted and the decay of human capital begins to overwhelm the network effects. Finally, the results are matched to dynamics in the late Roman empire, ancient China, and the modern West.

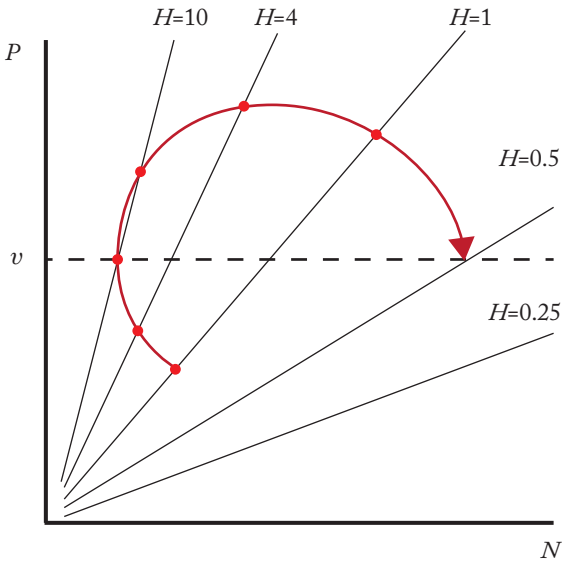


Fig 1. The rays from the origin indicate economies of scale of an individual's labor productivity P *in the local population* N*, with the marginal returns to scale a function of human capital* H*. Below the urbanization threshold when* P<υ*, the Malthusian rural environment selects for increasing human capital. When* P>υ*, the society urbanizes. The increased opportunity cost of intergenerational transmission of human capital then causes* H *to decline, even as* P *continues to rise. Eventually the decline in* H *overwhelms the productivity increases and results in deurbanization and a return to Malthusian conditions.*

(Poster 11 was canceled)

Detectability of Heterogeneous Communities in Networks using Matrix Eigenvectors

Bao Huynh[†,1] and Dane Taylor[†,2]

[†]Department of Mathematics, University at Buffalo, State University of New York
[1]baohuynh@buffalo.edu, [2]danet@buffalo.edu

Abstract

We present a method to estimate the eigenvalue distribution of graph-encoding matrices such as adjacency and Laplacian matrices. We study random graphs with N nodes generated by the stochastic block model (SBM). Our method predicts the asymptotic $N \to \infty$ behavior of eigenvalues and eigenvectors of graph-encoding matrices for SBMs with two planted communities. The planted communities are detectable if and only if there is a gap between the isolated and the bulk eigenvalues. We show that the detectability phase transition can be quantified by $\frac{\partial \lambda_2}{\partial \theta}$, where λ_2 is the second largest eigenvalue of the adjacency matrix and θ is proportional to the difference $p_{in} - p_{out}$ of inner p_{in} and inter p_{out} community edge probabilities. We also show that in the limit $N \to \infty$, $\frac{\partial \lambda_2}{\partial \theta}$ converges almost surely to $|\mathbf{v}^T \mathbf{w}|^2$, where $\mathbf{v}$ is the eigenvector corresponding to λ_2 and $\mathbf{w}$ is its expectation. Our method can be used to study the effects of community structure on Support Vector Machine (SVM) algorithm.

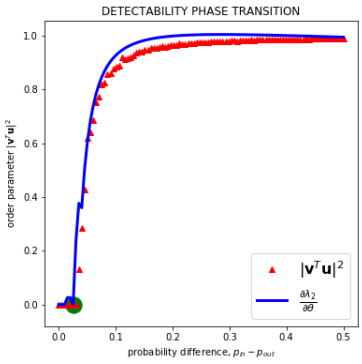


Figure 1: Detectability depends on the difference $p_{in} - p_{out}$, where p_{in} is the probability of an edge between two nodes in the same community and p_{out} is the probability of an edge between two nodes in different communities. The transition from undetectable (i.e.,$|\mathbf{v}^T\mathbf{w}|^2 = 0$) to detectable (i.e.,$|\mathbf{v}^T\mathbf{w}|^2 >> 0$) happens at the green dot when $pin - pout \simeq 0.02$.

(Poster 13 was canceled)

Spatiotemporal Dynamics of Housing Prices: A Bayesian Network Approach

S. Salinas-Ayaviri[1], *M.Villena*[1]. *(1) Adolfo Ibañez University, Santiago, Chile, silvia.salinas@edu.uai.cl.*

Housing markets play a crucial role in economies, the market value of properties is of great interest to local authorities, mortgage institutions, dissolved companies and other market participants. Following the 2007-2008 Global Financial Crisis, there have been growing research interests on the spatial interrelationships between house prices in many countries. Using Bayesian Graphical Vector Autoregressive (BG-VAR), we investigated the complex spatiotemporal dynamics of Chile housing prices (2002-2017). This network approach enables a data driven identification of the most dominant district where house price shocks may largely diffuse through the housing market and it is suitable for analyzing the complex spatial interactions between house prices. Using temporal house price volatilities for owner-occupied dwellings, the results show evidence of house price diffusion pattern in distinct sub-periods from different district housing submarkets in Santiago. We observed particularly prior to the crisis, diffusion of temporal house price volatilities. We found that during the evolution of Chile housing market, the prices diffuse in complex ways that require geographical clusters. The splitting and merging of clusters indicate that the house prices converge. Thus, we show that there are different classifications for converging clusters in different time periods.

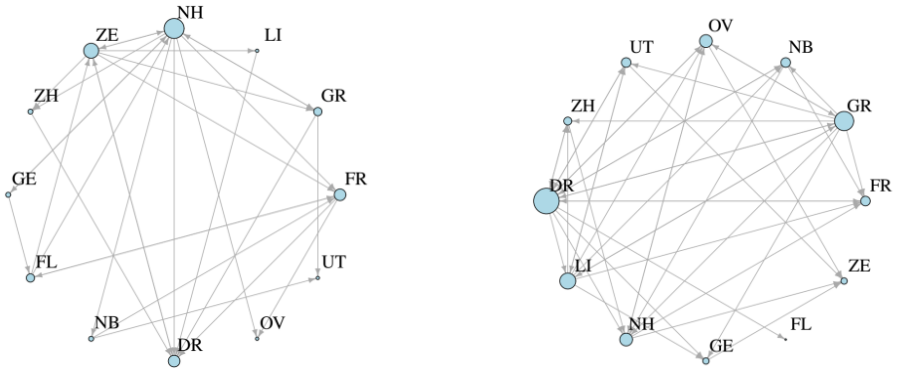


Fig.1. Network diagrams of temporal volatility connectedness in the regional market among the all districts during 2002 – 2008 (left figure), and 2008 – 2017(right figure). The size of the nodes is proportional to their degree (in-degree + out-degree).

Acknowledgements

SS-A is supported by the National Commission for Scientific and Technological Research (CONICYT) of Chile, Grant national doctoral scholarship. Also acknowledges to Center for Territorial Intelligence of Adolfo Ibañez University for the data bases.

Coupling of an agent-based model of HIV transmission dynamics with a mathematical model of intra-host dynamics of viral load

Python Ndekou Tandong Paul[1]

Ndekou William Françis[1]

Haby Camara[1]

[1]Cheikh anta Diop University, Department of Mathematics and Computer Science

pppython@yahoo.fr

In this paper, we have developed a nonlinear mathematical model describing the dynamics of the HIV evolution within-host taking into account the immune response. The mathematical model describes the dynamics of uninfected target cells, infected cells, and the dynamics of viruses. The immune target cells are CD4+ T cells susceptible to be infected by HIV viruses. This mathematical model is coupled with an intelligent agent-based model describing the transmission dynamics between individuals. The agent-based model allows to take into account social factors, population behaviors, education campaigns. Results showed that an infected individual can cause a high number of infections during the first 3 months corresponding to the period of the stimulation of antibodies, results also showed that education campaigns have reduced infections by at least 6%. We noted that, when the viral load is high, the number of CD4+ T cells decreases and it's possible to see the new symptoms of VIH. The probability of HIV transmission depends on the variations of viral load and the number of CD4+ T cells responsible for immune responses. We have shown that the infection status of the individual influences the spread of HIV within the community. The advantage of this coupled model comes from the fact that we take into account at the same time to the viral dynamics, the dynamics of immune cells, and the impact of social factors on the HIV transmission. The software platform that combines a non-linear mathematical model and an agent-based model provides a tool for predicting the spread of HIV in the communities.

Key words: **HIV prediction, agent based-model, mathematical model, coupled model, infection**

Heterogeneous Recurrence Analysis of Disease-altered Spatiotemporal Patterns in Multi-Channel Cardiac Signals

Ruimin Chen[1], Farhad Imani[1] and Hui Yang[1]

[1] Complex System Monitoring, Modeling and Analysis Laboratory
The Pennsylvania State University
rxc91@psu.edu

Abstract

Heart diseases alter the rhythmic behaviors of cardiac electrical activity. Variability analysis of successive heart beats is widely used for the detection of heart diseases. Recent advances in sensing technology bring 3-lead Vectorcardiogram (VCG), which presents rich information about space-time electrical activity of the heart. Traditionally, recurrence analysis of 12-lead electrocardiogram (ECG) signals is considered to investigate the recurrence behaviors in cardiac electrical activity. However, conventional recurrence analysis treats all the states homogeneously, whereas due to the high-level complexity in cardiac systems, there exist several different types of recurrence patterns in VCG signals. This paper presents the heterogeneous recurrence analysis of VCG signals for the identification of pertinent patterns about myocardial infarctions (MI) through multivariate monitoring and control. Few, if any, of previous approaches considered heterogeneous recurrence dynamics that underlie 3-lead VCG signals and used heterogeneous quantification analysis (HQA) for the identification of cardiac disorders. The heterogeneous recurrences quantifiers, as sensitive biomarkers, allows for the detection of myocardial injury through the VCG signals. The core components of the proposed methodology are shown in Figure. 1. Experimental results show that the proposed approach effectively characterize the heterogeneous recurrence patterns for the identification of disease-altered cardiac dynamics. The proposed heterogneous recurrence analysis of sensor signals for anomaly detection can be potentially extended into a variety of engineering domains.

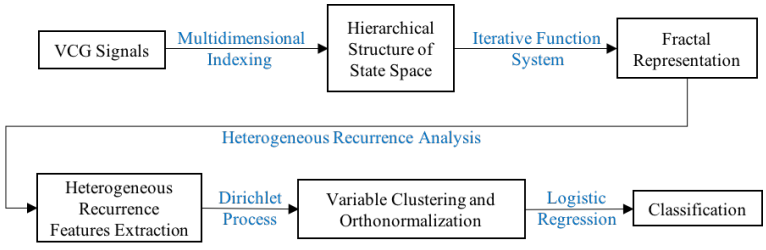


Figure 1: The flowchart of the heterogonous recurrence analysis of VCG signals

Markov Decision Process and Multifractal analysis for Image-Guided Additive Manufacturing

Farhad Imani[1], Bing Yao[1], Ruimin Chen[1], Prahalada Rao[2] and Hui Yang[1]

[1] Department of Industrial and Manufacturing Engineering,
The Pennsylvania State University, University Park, PA
[2] Department of Mechanical and Materials Engineering,
University of Nebraska, Lincoln, NE

Abstract

Additive manufacturing (AM) is a process to produce three-dimensional parts with complex and free-form geometries layer by layer from computer-aided-design models. However, real-time quality control is the main challenge that hampers the wide adoption of AM. Advancements in sensing systems facilitate AM monitoring and control. Realizing full potentials of sensing data for AM quality control depends to a great extent on effective analytical methods and tools that will handle complicated imaging data, and extract pertinent information about defect conditions and process dynamics plot variable. This paper considers the optimal control problem for AM parts whose layerwise defect states can be monitored using advanced sensing systems. Specifically, we formulate the in situ AM control problem as a Markov decision process and utilize the layerwise imaging data to find an optimal control policy. We take into account the stochastic uncertainty in the variations of layerwise defects and aim at mitigating the defects before they reach the nonrecoverable stage. Finally, the model is used to derive an optimal control policy by utilizing the defect-state signals estimated from layerwise images in a metal AM application.

A game theoretic approach to modeling dynamics of Amyloid-β aggregation along competing pathways

Joseph Pateras, Edward Steen, and Ashwin Vaidya

Aggregation of amyloid β (Aβ) peptides is a significant event that underpins neurodegenerative diseases such as Alzheimer disease (AD). Aβ aggregates, especially the low-molecular weight oligomers, are the primary toxic agents in AD. Therefore there is increasing interest in understanding their formation and behavior. Aggregation is a nucleation-dependent process in which the pre-nucleation events are dominated by Aβ homotypic interactions making this a very complex system. In this poster we will discuss our investigations of heterotypic interactions between Aβ and fatty acids (*FA*s) via mathematical modeling and game theoretic tools undertaken by our research group.

We observe that *FA*s influence aggregation dynamics in three broadly-defined *FA* concentration regimes containing non-micellar, pseudomicellar or micellar phases. While the non-micellar phase promotes on-pathway fibrils, pseudomicellar and micellar phases promote predominantly off-pathway oligomers. Off-pathway oligomers saturate within a limited molecular size, and likely with a different overall conformation than those formed along the on-pathway, suggesting the generation of distinct *conformeric strains* of Aβ, which may have profound significance for our understanding of the biophysics of protein aggregation.

The model involves a coupled system of N nonlinear differential equations which model the aggregation behavior of the Aβ molecules along the two main pathways discussed above, which are linked through intermediate bridges. Here, N refers to the dimensionality of the model which is taken to be either 4 or 6. Our task is therefore to investigate the flow and evolution of these amyloid-proteins in a very complex network and find conditions under which each of the pathways dominate. Results of our computations validate previous experimental observations and provide insights into potential influence of biological interfaces in modulating protein aggregation pathways.

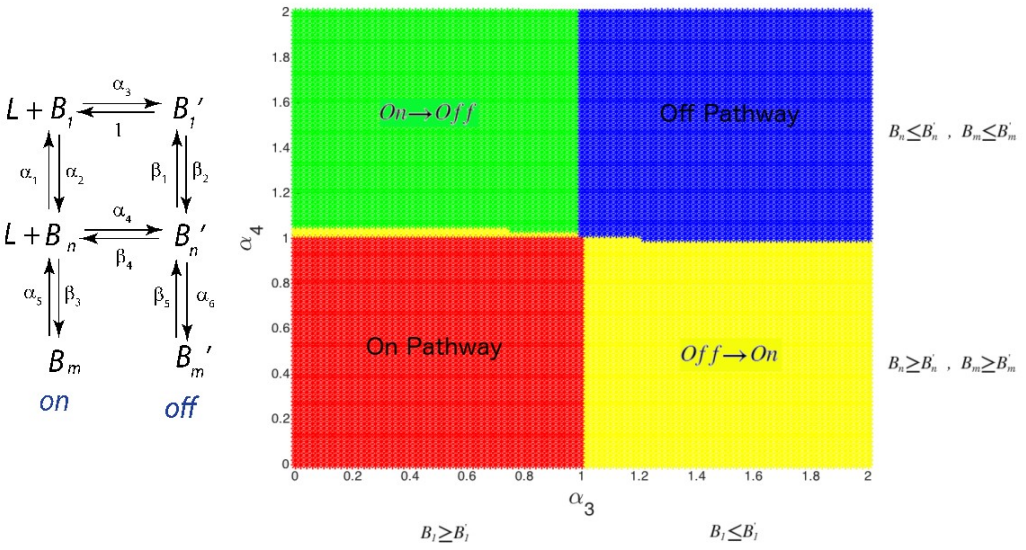


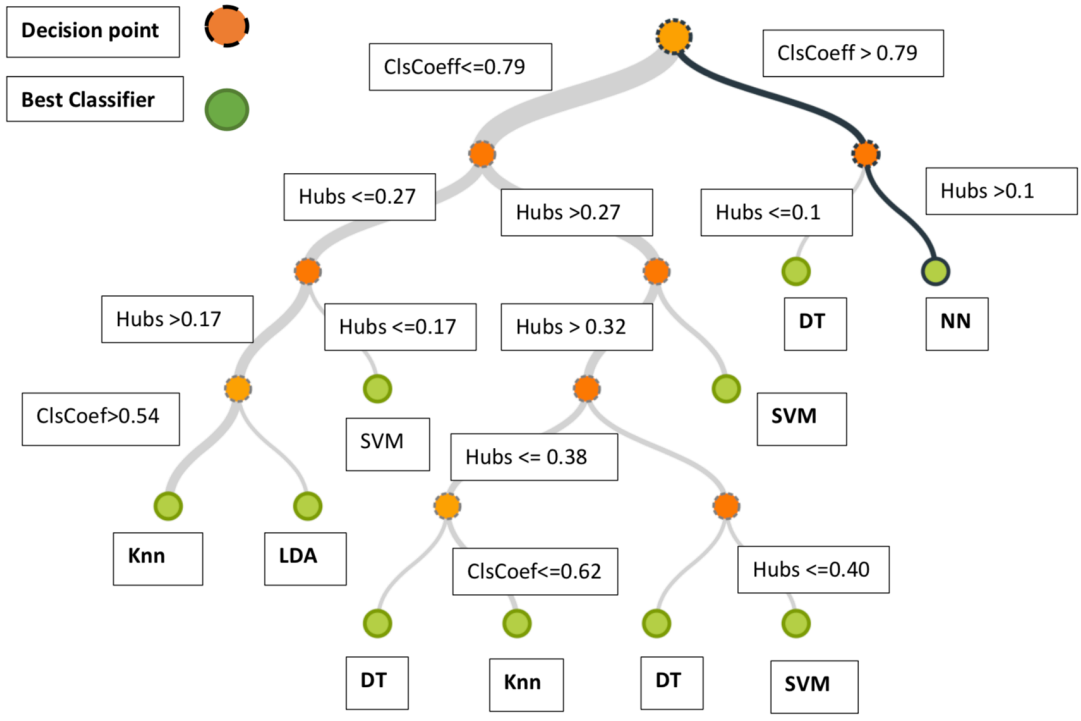
Figure: This figure on the left shows a schematic of the biophysical reaction leading to various pathways of A-beta aggregation. The figure on the right depicts a phase diagram showing the competing pathways under the right conditions, depending on the magnitudes of two 'bridge' reaction rates.

Importance of Network Metrics in Classification

Dieudonne Ouedraogo

Abstract

In machine learning, the performance of a classifier is intrinsically related to the task. The structure between data points within the dataset plays an important role. In this paper, we explore the usage of network metrics to describe the selection of a machine learning algorithm for a classification task concerning a specific dataset. A dataset is transformed into a graph representation based on the ϵNN algorithm. A data point is a node, and an edge exists between two points i,j if $d(i,j) < \epsilon$. A post-processing step is applied to the graph, pruning edges between examples of different classes. The structural information such as density, clustering coefficient, and hubs are extracted. Various data sets are collected, their network metrics are computed. A predictive model is built to investigate the possible relationship between networks characteristics and the classifier used to be used on the machine learning task. Results show that network metrics such as clustering coefficient, hubs,density are very informative in predicting the classifier to be used on the task. For example, clustering coefficient greater than 0.79 or hubs greater 0.1, neural network are the best performers; for hubs less than 0.1, using decision trees gives the highest accuracy on the predictions.



Decision Tree with network metrics as variables and algorithms as outputs

Countries' Diversification and Transition to Green Economy

Seyyedmilad Talebzadehhosseini[1,2], Chathika Gunaratne[1,3], Steven Scheinert[1,4], and Ivan Garibay[1,5]

[1] Complex Adaptive System Laboratory
Department of Industrial Engineering and Management Systems
University of Central Florida, Orlando, Florida
[2]talebzadeh@knights.ucf.edu,[3]Chathika.Gunaratne@ucf.edu,[4]steven.scheinert@ucf.edu,[5]igaribay@ucf.edu

As the usage of natural resources increases due to the increase of world population, the economic, social, and environmental costs of traditional production will increase, creating more pressure to change to address these costs. A transition to a green economy requires the transformation of a nations production capabilities, infrastructure, technology, labor, and capital to more green modes of production and utilization. Existing research has examined the potential for future expansion (Mealy and Teytelboym, 2018) and has examined the paths that countries take for expansion of non-green product capabilities (Hidalgo et al., 2007) but not how countries have expanded their green production capabilities. That research developed the product space (Hidalgo et al., 2007) and green product space concept (Mealy and Teytelboym, 2018) to examine that transition of production and economic diversification. This research applies that concept to study the paths of diversification that different economies have followed. Figure 1 depicts twelve green products according to (Hidalgo et al., 2007) method for identifying a product space. The products are linked by the similarity in their capabilities they share, shown using green links. Closely located products are similar and so represent a more likely candidate for innovation from non-green to green production.

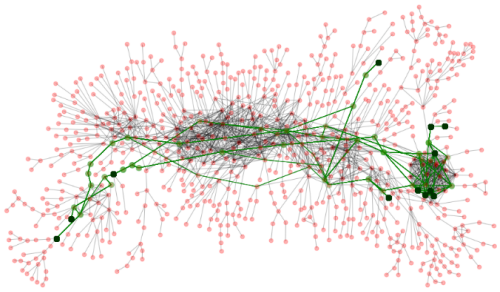


Figure 1: Shortest paths between all green products of the product space.

This research uses complex systems approach to develop an agent-based simulation model to explore empirical pathways of economy for green production development and to examine the pattern of producing new green products based on the products that economy is already producing. The results of this research will show how economies transitioned to green production and what pattern for producing new green products they followed.

The network of occupation space needs for economic improvement

Mostafa Saeidi[1], Ramya Akula[1], Steven Scheinert[1], Anamaria Berea[1] and Ivan Garibay[1]

[1] University of Central Florida, Complex Adaptive Systems Laboratory
msaeidi@knights.ucf.edu, ramya.akula@knights.ucf.edu, steven.scheinert@ucf.edu
anamaria.berea@ucf.edu, igaribay@ucf.edu

Abstract

As skills diversify the occupational structure of each region, the connections between skills demonstrate the occupation diversification. Existing research has investigated the occupation diversification according to the skills network (Muneepeerakul et al., 2013 Anderson, 2017; Yildirim and Coscia, 2014) and the location of U.S metropolitan areas in the skills network (Alabdulkareem et al., 2018) , but not how the location of cities in the network has changed and how occupation diversification occurs in an urban area over time. Understanding occupations diversification is relevant since it drives diversity in the local economy which drives the health of the local economy and diversity of the labor market. This research applies existing skills network which was introduced by Alabdulkareem et al. (2018) to study occupation diversification of metropolitan areas over time and investigate the pattern of city expansion and contraction in the skills space network. The results of this research will present a new form of economic structure based on the occupation diversification which helps policymakers to increase innovation in the service sector, reduce the unemployment rate, and introduce new occupations that are related to the previous occupations in the region. To conduct this research, a quantitative comparison is presented. Applying the skills data for six U.S. metropolitan areas in the 10-year period and visualizing where metropolitan areas are located in the skills space network over time. Figure 1 shows the density of relationships among occupations in different types of occupations. This research applies a data-driven methodology and network science to map skills space for selected metropolitan areas and explore how metropolitans position in the network change over time.

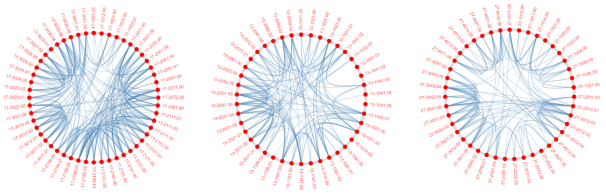


Figure 1: The density of interactions between occupation in three different service sectors (Architecture and engineering - Business and financial operations - Arts, design and media) $[0, 10]$.

(Poster 24 was canceled)

A pharmacokinetic model of lead-calcium interactions

Steven Lundgren[1], Anca Rădulescu[2]

[1]Mathematics, State University of New York, New Paltz

[2]Mechanical Engineering, State University of New York, New Paltz

lundgres2@hawkmail.newpaltz.edu

Lead is a naturally-occurring element. It has been known to man for a long time, and it is one of the longest established poisons. The current consensus is that no level of lead exposure should be deemed "safe." New evidence regarding the blood levels at which morbidities occur has prompted the CDC to reduce the screening guideline of 10 μg/dl to 2 μg/dl. Measurable cognitive decline (reduced IQ, academic deficits) have been found to occur at levels below 10mg/dl.

Knowledge of lead pharmacology allows us to better understand its absorption and metabolization, mechanisms that produce its medical consequences. Based upon an original and very simplified compartmental model of Rabinowitz (1973) with only three major compartments (blood, bone and soft tissue), extensive biophysical models sprouted over the following two decades. However, none of these models have been specifically designed to use new knowledge of lead molecular dynamics to understand its deleterious effects on the brain. We build and analyze a compartmental model of lead pharmacokinetics, focused specifically on addressing neurotoxicity. We use traditional phase space methods, parameter sensitivity analysis and bifurcation theory to study the transitions in the system's behavior in response to various physiological parameters.

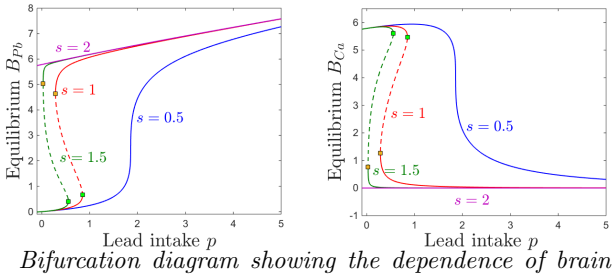


Fig: *Bifurcation diagram showing the dependence of brain lead and calcium steady states on lead intake, for various efficiency modes of the blood brain barrier.*

We conclude that modeling the complex interaction of lead and calcium along their dynamic trajectory may successfully explain counter-intuitive effects on systemic function and neural behavior which could not be addressed by existing linear models. Our results encourage further efforts towards using nonlinear phenomenology in conjunction with empirically driven system parameters, to obtain a biophysical model able to provide clinical assessments and predictions.

Template iterations of quadratic maps and hybrid Mandelbrot sets

Anca Rădulescu, Brandee Williams, Kelsey Butera

Mathematics, State University of New York, New Paltz

radulesa@newpaltz.edu

As a particular problem within the field of non-autonomous discrete systems, we consider iterations of two quadratic maps $f_{c_0} = z^2 + c_0$ and $f_{c_1} = z^2 + c_1$, according to a prescribed binary sequence (template). We study the parameter locus for which critical orbits are bounded (the "Mandelbrot set"). For a fixed template, one may consider this locus as a subset of $(c_0, c_1) \in \mathbb{C}^2$; for fixed quadratic parameters, one may consider the subset of templates which produce a bounded critical orbit. We consider both situations, as well as "hybrids," and study topological properties of these sets. We find that it is possible to tell which specific errors are more likely to affect the system's dynamics, in absence of prior knowledge of their timing. Moreover, within an optimal locus for the correct function, almost no errors can affect the sustainability of the system. Mathematically, our work complements broader existing results in random dynamics with more specific detail for the case of two random iterated functions, which is a valuable context for applications.

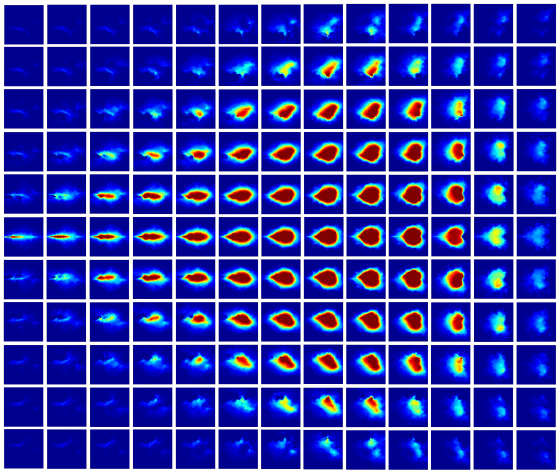


Figure: *Grid of hybrid Mandelbrod sets, illustrating the "error" locus for which the system viability (bounded critical orbit) depends on the timing of the errors.*

We discuss the potential of using this mathematical framework (in which one can apply a "good" transformation f_{c_1} versus an "erroneous" f_{c_0}) to study the effect of errors in copying mechanisms (such as DNA replication). We consider problems that a sustainable replication system may have to solve when facing the potential for errors. We find that it is possible to tell which specific errors are more likely to affect the systems dynamics. We find that it is possible to tell which specific errors are more likely to affect the systems dynamics, in absence of prior knowledge of their timing.

Colonel Blotto Games for Cyber-physical Systems Security: Cooperative Games

Victoria Kim[1] Matthew Dunn[1] and Emrah Akyol[1]

[1] Department of Electrical and Computer Engineering, Binghamton University, SUNY
{vkim5, mdunn8, eakyol}@binghamton.edu

The "Colonel Blotto" Game was introduced by Borel and Ville in 1921 [1]. The game is defined as two opposing generals battling across several battlefields in which each general allocates resources to each battlefield, simultaneously. The general who places the greater number of resources at a battlefield wins that battlefield, and the general that wins more battles, wins the game.

In the previous work [2], each attacker is modeled as an individual, uncooperative agent. In this paper, building on prior work [2], we analyze coalitional games within the same context. Our objective is to numerically analyze and determine what type of game system parameters will cause individual attackers to form stable coalitions with each other.

Our model consists of three attackers (A, B, C) and one defender who each have unique goals based on their value the physical nodes. Attacker 1 values Physical Node 1 four times more than Physical Node 2. Attacker 2 values Physical Node 2 four times more than Physical Node 1. Attacker 3 and the defender value each physical node equally.

The attackers can either form no coalition, a sub-coalition or the grand coalition. Each case is treated as a non-cooperative game between the coalition of attackers, any attacker acting alone and the defender. Each coalition is treated as a singular player whose resources are the combined resources of all players within that coalition. The total payoff given to a coalition is based on how much each individual attacker values the physical nodes. The coalition's value of a physical node is represented as the average of each individual member's value of that physical node. We assume that the total payoff given to each party is then divided between each individual player using the Shapley value formula. Numerical analysis shows that it is beneficial for Attacker 3 to form any coalition, especially the grand coalition. Conversely, the grand coalition provides the lowest payoff for both Attacker 1 and Attacker 2. Thus, both Coalition 1 (no coalition) and Coalition 3 are unstable imputations. Coalition 2a is unstable because Attacker 3 is incentivized to join BC, creating the grand coalition. Coalition 2c is the only stable imputation because no player is motivated to deviate and provides players with a higher payoff than Coalition 2b.

Payoff Matrix					
		Attackers			Defender
No Coalition (1)	{A}, {B}, {C}	1.2346	1.2067	0.9774	0.0448
Partial Coalition (2a)	{AB}, {C}	2.5006		0.9999	0.0084
Partial Coalition (2b)	{BC}, {A}	2.2589		1.2476	0.0406
Partial Coalition (2c)	{AC}, {B}	2.2507		1.2500	0.0084
Grand Coalition	{ABC}	3.5044			0.0548

Table 1: Table 1: Payoff Matrix

Payoff Matrix: Shapley Value					
		Attacker 1 {A}	Attacker 2 {B}	Attacker 3 {C}	Defender
No Coalition (1)	{A}, {B}, {C}	1.2346	1.2067	0.9774	0.0448
Partial Coalition (2a)	{AB}, {C}	1.2643	1.2364	0.9999	0.0084
Partial Coalition (2b)	{BC}, {A}	1.2476	1.2441	1.0148	0.0406
Partial Coalition (2c)	{AC}, {B}	1.2580	1.2500	1.0009	0.0084
Grand Coalition (3)	{ABC}	1.1327	1.1187	1.2529	0.0548

Table 2: Table 2: Payoff Matrix: Shapley Value

References

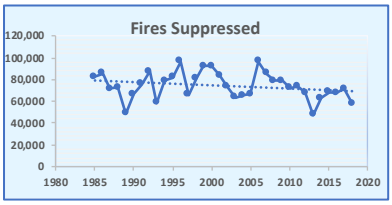
[1] E. Borel, J. Ville, "Applications de la theorie des probabilites aux jeux de hasard." J. Gabay, 1938.

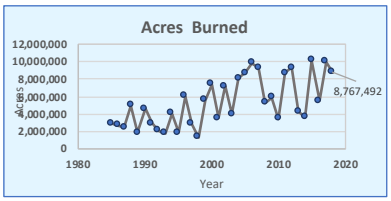
[2] J. Headrick, J. Liang, E. Akyol, "On the impact of network connectivity in Colonel Blotto games." NERCCS, 2018.

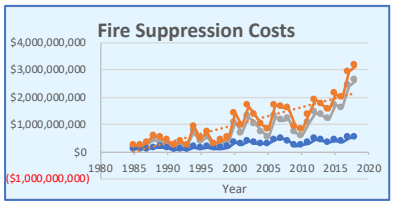
Wildfire regimes – research in complex systems science

Lorren Kay
Center for collective Dynamics of Complex Systems
Binghamton University, State University of New York
lkay1@binghamton.edu

A wildfire regime is the pattern, frequency, and intensity of fire established in an area over a long period of time. The patterns in a particular ecosystem arise from a complex set of interactions between land, vegetation, climate and ignition sources. Over the past twenty years wildfire regimes have been dramatically shifting. Fires are becoming more intense and frequent with longer fire seasons affecting nearly every continent on Earth. Empirical evidence indicates two primary factors — climate change and a highly successful fire suppression policy instituted in the 20th century. In 2015 the United States reportedly extinguished ninety-seven percent of all fires before reaching 120 ha, with a cost ranging in the billions of dollars.







The increasing fire suppression costs are unsustainable. Additionally, research indicates fire adapted forests rely on cleansing fires for long term health, calling into question the fire suppression policy. The risk this policy poses to forest health is exemplified in the western region of the United States. Forests in this region have coevolved with fires but have been heavily suppressed yet saw one of the worst fire seasons in history. More than 8,000 fires were responsible for burning land in excess of 1.8 million acres (766,439 ha).

New models are required to address the wildfire threat. Wildfire research should benefit from complex systems science however, fire and disaster scientists are disappointed with complex systems theory. Criticisms cite a lack of theoretical basis for many of the empirical studies related to emergence and self-organization with current theoretical explanations developed by scientists in other disciplines. Complex systems science must develop more theoretical and applied level work, showing its applicability and usefulness as a "lens" to understand, examine, plan for, and respond to wildfires and other disasters. [1] This paper lays a foundation to work towards this goal by surveying the current landscape of principles and methodologies. Future research shall build upon this foundation to develop a synthesis of complex systems science theories and applied research needed by scientists searching for new methods to create resilient and sustainable ecosystems and better predictive capabilities for fire prone ecosystems.

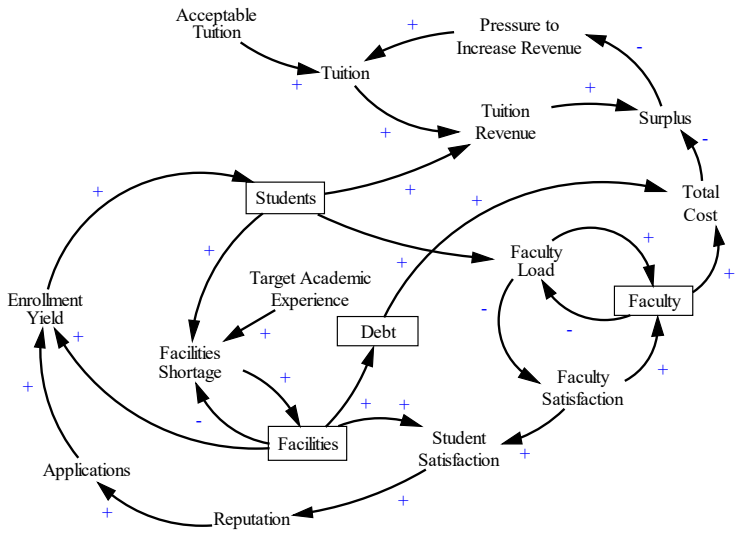
References

1. Etkin, D. (2008). *Disaster Theory*. Elsevier Inc.

Dynamics of Complex Service Systems

Oleg Pavlov, Worcester Polytechnic Institute
Evangelos Katsamakas, Fordham University

We suggest that important trends in education, such as escalating tuition, can be understood if we view and study universities as complex systems that provide educational services. Methodologically, we draw on the service science theoretical framework [1], which we adapt to education [2], and then supplement our analysis with a computational model. The fundamental concept of service science is the service system, which is a dynamic network of resources that produce and deliver value to stakeholders. Examples of service systems are numerous and include firms, towns, government agencies, and even countries.



We build a system dynamics [3] model that represents a typical US university. As shown in the causal diagram, the model takes into account complex interrelationships between faculty, physical plant, tuition revenue, financials, reputation and service outcomes. Variables that accumulate, and therefore introduce delays, are shown as rectangles. Plusses and minuses indicate positive and negative causal relationships. The model has five sectors: Students, Faculty, Facilities, Financials and Measurements.

We use the model to analyze the strategic behavior and resulting dynamics of a typical tuition-based university in the context of rapid and accelerating transformation caused by market forces. We identify and compare policies that are important to various stakeholders. We also stress-test the university against negative shocks, such as an unexpectedly high enrollment yield or a dwindling population of international students, who typically do not receive financial aid.

References:

1. Spohrer, J., et al., *Steps Toward a Science of Service Systems.* Computer, 2007. **40**(1): p. 71-77.
2. Pavlov, O. and F. Hoy, *Toward the Service Science of Education*, in *Handbook of Service Science, Volume 2*, P.P. Maglio, et al., Editors. 2019, Springer: New York, NY.
3. Morecroft, J., *Strategic Modelling and Business Dynamics: A Feedback Systems Approach*. 2007: Wiley, John & Sons.

Emotive Impact on Facebook, in the presidential campaign of Mexico 2018

Carlos Augusto Jiménez Zarate
uanl.jimenez@gmail.com

Abstract

This research will analyze the impact of the emotions emitted in the Facebook fan pages of the candidates in the recent presidential elections in Mexico. The current electoral campaigns are supported by the dissemination of content or propaganda in order to generate support in the electoral preferences. Within the results of the present investigation, a correlation is highlighted between the emotions generated by the publications of the candidates and the electoral result in the recent presidential election in Mexico. The results of the research are an opportunity to understand the emotional effect of political propaganda, in addition that can be useful for the construction of a model of electoral prediction, based on the emotions generated in Facebook.

Keywords: Sentiment on Facebook, Emotive impact, electoral preferences and linear correlation.

1. Introduction

One of the most popular measures of "success" of a publication in any socio-digital environment, is the total reach Organic and non-organic,[1] this measurement represents the number of users who saw, interacted directly or indirectly with any publication. But the information of scope, is not enough to have a real approximation of the success or effectiveness of a publication. To complement the information, other factors that are proposed in this research must be taken into account, such as the feeling and the emotional reactions that the publications can generate. It has been shown that the emotions or emojis emitted in Facebook, have a correlation in the detection of the feeling of the users [2].

2. Method

This investigation has a descriptive scope, since it describes the situations, the contexts and the information of how the users will manifest through the use of emojis on Facebook, and correlational because a correlation analysis will be made between the proposed metrics and the electoral preferences. (Sampieri, 2014).

The "likes" emitted in the analyzed fanapges will be excluded in this investigation because they can be influenced by artificial positioning systems (astroturfing). The information collected were databases of public access on Facebook. We used the "Fanpage Karma" monitoring tool with which we obtained the total of published publications, their reactions, the number of followers and followers won, these measurements were made directly, from the official pages of the candidates. The analysis period was from December 14, 2017 to June 27, 2018. The method of analysis proposed in this investigation is condensed in figure 1.

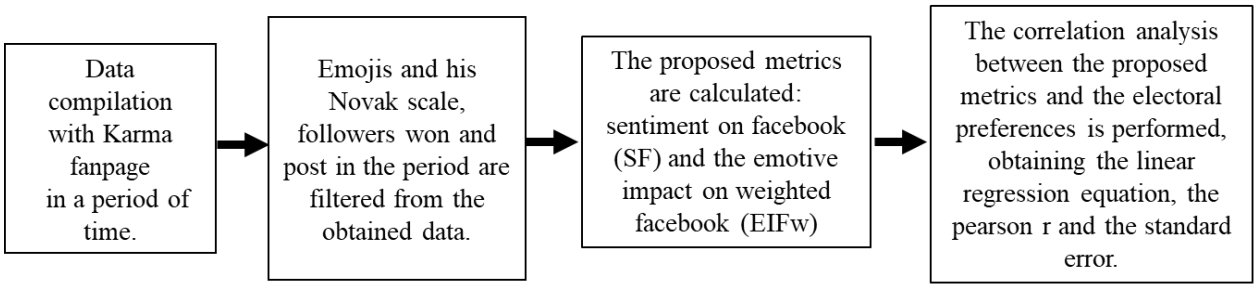


Figure 1. Proposed method of Emotive Impact on Facebook (EIF).

3. Purpose

The objective of this research is to establish the emotive impact on the official Facebook fanpage of the candidates during the Mexico 2018 presidential election process in the electoral preferences.

4. Data Collection

The federal electoral process, for the presidency of Mexico, was subdivided into phases or periods, according to the official calendar of the National Electoral Institute, the periods or electoral phases, for the presidency of Mexico, were limited by the following dates [4] :

- Pre-Campaign from December 14, 2017 to February 11, 2018 (60 days).
- Inter-Campaign from February 12 to March 29, 2018 (45 days)
- Campaign from March 30 to June 27, 2018 (90 days)

The day of the election was the first of July 2018.
The candidates registered for the presidential election were:
Ricardo Anaya Cortés or “ANAYA” (nickname), registered by the "For Mexico to the Front" coalition composed of the following parties: National Action Party “PAN” (acronym in Spanish), Democratic Revolution Party “PRD” (acronym in Spanish) and Movimiento Ciudadano “MC” (acronym in Spanish). Jose Antonio Meade Kuribreña or “MEADE” (nickname) registered by the coalition "All for Mexico" composed of the following parties: Institutional Revolutionary Party “PRI” (acronym in Spanish), Green Party “VERDE” (acronym in Spanish) and the New Alliance party. Andrés Manuel López Obrador or “AMLO” (nickname) registered by the collation "Together We Will Make History", integrated by the following parties: Movement of National Regeneration “MORENA” (acronym in Spanish), Labor Party “PT” (acronym in Spanish) and Social Encounter Party “PES” (acronym in Spanish). Jaime Heliodoro Rodríguez Calderón or “BRONCO” (nickname) registered as an Independent candidate. Margarita Ester Zavala Gómez del Campo registered as an Independent, but the data generated by her publications and interactions, is not included in this research because she resigned her candidacy on May 16, 2018.

Facebook has positioned itself as the most important social network at the moment, in a short time it managed to extend its user outside the university communities, passing to a generalized use in the general population [5]. A fan page on Facebook is a website where you can make publications or publications, of different types of content and followers or users of Facebook can react or emit emotions through the emoticons. It is estimated that only about 8% of followers of a fan page see publications and 80% of Facebook users ignore fanpage publications [6].

The action of sharing a publication is more likely to occur to a large extent when publications offer information, or are based on a sensory propaganda strategy, this may mean that for a user, sharing a publication has more to do with the affective and or cognitive [7]. As of February 2016, Facebook gives users and followers of fan pages the opportunity to express feelings through emoticons or emojis in publications. This type of emotional reaction can offer us a greater understanding of the social feeling that a publication can generate. In the research "Sentiment emojis" (Novak, 2016) a score is defined for each emoji, on a scale of a spectrum of feelings, which ranges from -1 to +1. Where the emoji score that approaches -1 is more negative, and the emoji score that approaches +1 is more positive. Table 1 shows the number of emoticons, which were generated by the publications in each period of the electoral process, in addition to the weighting factors of the feeling of each emoji, the highest value in Facebook is the emoji of "I love" with a value of 0.746 and the emoji with the lowest value is "I get angry" with a value of -0.173.

Table 1. Emojis generated by the publications of the candidates by period, of the presidential electoral process in Mexico 2018.

	Emoticons on Facebook in Pre-Campaign (December 14, 2017 to February 11, 2018)					
Candidates (Nicknames)	Love (0.746)	Haha (0.409)	Wow (0.269)	Sad (0.007)	Angry (-0.173)	Total Emojis
ANAYA	127,200	266,312	5,291	2,143	47,382	448,328
MEADE	90,718	128,895	3,783	1,604	35,217	260,217
AMLO	270,285	86,659	18,509	850	3,364	379,667
BRONCO	68,606	78,736	4,887	1,345	14,312	167,886
	Emoticons on Facebook in Inter-Campaign (February 12 to March 29, 2018)					
Candidates (Nicknames)	Love (0.746)	Haha (0.409)	Wow (0.269)	Sad (0.007)	Angry (-0.173)	Total Emojis
ANAYA	54,788	62,756	7,016	781	10,954	136,295
MEADE	59,824	57,647	2,236	605	15,784	136,096
AMLO	196,425	36,729	10,812	460	1,799	246,225
BRONCO	75,575	74,103	6,356	2,970	9,833	168,837
	Emoticons on Facebook in Campaign (March 30 to June 27, 2018)					
Candidates (Nicknames)	Love (0.746)	Haha (0.409)	Wow (0.269)	Sad (0.007)	Angry (-0.173)	Total Emojis
ANAYA	1,058,438	2,876,016	64,306	24,379	184,272	4,207,411
MEADE	459,134	892,931	18,168	6,711	111,866	1,488,810
AMLO	1,288,193	66,831	85,287	2,450	9,669	1,452,430
BRONCO	971,976	686,641	35,118	14,540	39,829	1,748,104

Source: Data extracted from the official fanpages of the presidential candidates in Mexico, through the analysis tool "FanpageKarma.com", during the period from December 14, 2017 to June 27, 2018, the values of the emojis correspond to the scale of Novak

5. Procedure

In the research "Facebook sentiment: reactions and emojis" (Tian et al., 2017), he proposes the measurement of feeling, based on emojis in the comments. For this research we propose the formula "Sentiment on Facebook" (SF), for the measurement of sentiment based on the emojis generated in the publications on Facebook in a period of time. The higher the value of sentiment in Facebook, it means that the emotional interaction in the analyzed fanpage is more positive.

$$Sentiment\ on\ facebook = \frac{[(\sum Love*0.746)+(\sum Haha*0.409)+(\sum Wow*0.269)+(\sum Sad*0.007)-(\sum Angry*0.173)]}{\sum Emotional\ reactions\ of\ the\ fanpage\ in\ the\ period} \quad (1)$$

The "Emotive Impact on Facebook" (EIF), in this research is proposed to calculate by multiplying the "Followers Won" (FW) in the period analyzed and the "Sentiment on Facebook" (SF), then divide the result of the factors by the number of publications in the period. The higher the value is the emotive impact on Facebook, it means that the social interaction in the analyzed fanpage, is more powerful and efficient with respect to the number of posts, in addition to having a more positive sentiment.

$$Emotive\ Impact\ on\ Facebook\ (EIF) = \frac{(FW)(SF)}{Posts\ in\ the\ period} \quad (2)$$

6, Results

Table 2 shows the followers won, the posts in the period, the sentiment generated by each fanpage of the candidates, as well as the emotive impact metric on Facebook and finally the emotional impact weighted, which serves as the basis for the competitive comparison between the candidates. The results in each electoral period show that it was AMLO the candidate who obtained the greatest emotive impact on Facebook.

Three aspects are relevant: First, the number of followers generated by the candidates, was the independent candidate BRONCO who obtained the largest number of followers in two periods, although this activity was not reflected in the electoral preferences, since he was the candidate with less electoral preference in the periods analyzed. Secondly, the number of posts that each candidate issued on their fanpage, in the three periods was again the independent candidate "BRONCO" with the largest number of posts issued from his fanpage, but his electoral preference never moved from the last place, however the candidate "AMLO" with fewer posts achieved a greater emotional impact. In third aspect, the measurement of sentiment on facebook, shows us how AMLO obtained the highest scores in the three periods of the campaign electoral process.

The three previous aspects give way to calculate emotive impact on facebook, where the measurements indicate how "AMLO" was the overwhelming winner throughout the electoral process, as well as in the weighted measurement of the emotive impact.

Table 2. Condensed of the results of sentiment and emotive impact of the publications of the candidates by period of the presidential electoral process in Mexico 2018.

Period	Candidates (Nicknames)	Followers Won (FW)	Posts	Sentiment on Facebook (SF)	Emotive Impact on Facebook (EIF)	Emotive Impact on Facebook weighted (EIFw)
Pre-Campaign (December 14, 2017 to February 11, 2018)	ANAYA	227,366	91	0.44	1,098	0.28
	MEADE	91,477	315	0.44	129	0.03
	AMLO	281,397	72	0.64	2,486	0.63
	BRONCO	85,245	420	0.49	99	0.03
Inter-Campaign (February 12 to March 29, 2018)	ANAYA	49,797	64	0.49	380	0.06
	MEADE	30,888	175	0.49	86	0.01
	AMLO	133,359	16	0.67	5,557	0.86
	BRONCO	243,772	328	0.51	382	0.06
Campaign (March 30 to June 27, 2018)	ANAYA	1,025,445	292	0.46	1,629	0.26
	MEADE	466,636	682	0.47	319	0.05
	AMLO	826,912	168	0.70	3,421	0.55
	BRONCO	1,248,188	1,160	0.58	621	0.10

Source: With the data extracted from the official fanpages of the presidential candidates in Mexico, through the analysis tool "FanpageKarma.com".

Table 3 shows the values obtained from the weighted emotive impact, which each candidate generated on his official Facebook page and the data on effective electoral preferences, that the electoral polling company Consulta Mitofsky [9] reported during the months of February, March and June, which correspond to the months of each period of the electoral process in Mexico. The emotive impacto n Facebook weighted (EIFw) was obtained by dividing the impact value of each candidate by the sum of the values of the impact of all the candidates.

The results in each period of the electoral process for the presidency of Mexico in 2018, shows that Andrés Manuel López Obrador "AMLO" the candidate who obtained the highest scores of the emotional impact on Facebook. Other relevant data of the measurement of the emotional impact are:

a) The variability of the candidate ANAYA, who in the first period obtained an index of 0.28 in the first period, while in the second period he obtained 0.06 and for the third or last period returned to raise his emotive impact index to 0.26

b) The candidate MEADE obtained very low scores of emotive impact on Facebook in the three periods.

c) The independent candidate BRONCO obtained a slight increase in the emotive impact during the three periods, but he was always far below the AMLO candidate.

Table 3. Emotive Impact on Facebook and electoral preference of candidates, by period of the presidential electoral process, Mexico 2018.

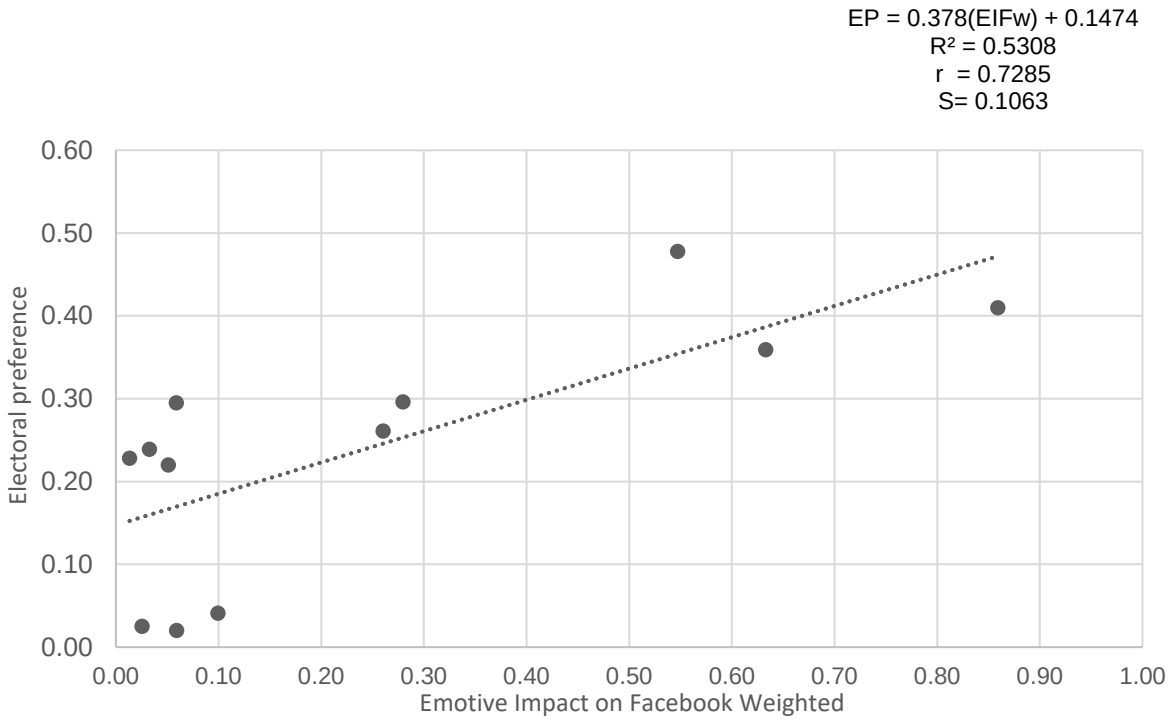
	Candidates (Nicknames)	Emotive Impact on Facebook weighted (EIFw)	Electoral Preference (EP)
Pre-Campaign (December 14, 2017 to February 11, 2018)	ANAYA	0.28	0.30
	MEADE	0.03	0.24
	AMLO	0.63	0.36
	BRONCO	0.03	0.03
Inter-Campaign (February 12 to March 29, 2018)	ANAYA	0.06	0.30
	MEADE	0.01	0.23
	AMLO	0.86	0.41
	BRONCO	0.06	0.02
Campaign (March 30 to June 27, 2018)	ANAYA	0.26	0.26
	MEADE	0.05	0.22
	AMLO	0.55	0.48
	BRONCO	0.10	0.04

Source: Data extracted from the official fanpages of the presidential candidates in Mexico, through the analysis tool "FanpageKarma.com", during the period from December 14, 2017 to June 27, 2018. The data of the electoral preferences they were extracted from the Mitofsky Consultation at the following website: https://elecciones.sociales.unam.mx/wp-content/uploads/2018/06/Mitofsky_Preferencias2018-Junio.pdf

To test the effectiveness of the "emotive impact" metric on Facebook, a correlation was made shown in graph 1, between the percentages of the Electoral Preferences, and the percentages of the emotional impact of the candidates per period.

The correlation of the obtained equation has a Pearson r = 0.7285 and a r^2 = 0.5308, with a standard error = 0.1063, that indicates that there is a deviation of the estimate of more than 10%, which for cases of correlation between variables is not as accurate for a prediction model. However, the result brings us closer to the compression of the emotive impact on Facebook and the electoral preferences, the results are shown in graph 1.

Graph 1. Correlation between the Emotive Impact on Facebook fanpages of the candidates and the Electoral Preference in the presidential electoral process, Mexico 2018.



Source: Data extracted from the official pages of the presidential candidates in Mexico, through the analysis tool "FanpageKarma.com", during the period between December 14, 2017 and June 27, 2018 and correlated with the electoral preferences of acerdo to the company Consulta Mitfosky. (https://elecciones.sociales.unam.mx/wp-content/uploads/2018/06/Mitofsky_Preferencias2018-Junio.pdf)

7. Conclusion

The results indicate that a large number of publications or posts is not significant to generate a greater propagandistic impact, on the contrary they are a burden for the measurement of effectiveness of the impact on emotions that can be expressed by emojis on Facebook. The results of the measurement of the "sentiment" generated by the emojis and its scale Novak, showed that the winning candidate of the election was the one that more obtained in the three periods analyzed the highest scores. The contents emitted from the official Facebook fanpages of the presidential candidates in Mexico in 2018, generated an emotional impact that correlated positively with the electoral preferences in each period of the electoral process, although the value of the standard error of the correlation limit the results.

This research, can be taken as a basis for the construction of a predictive model of electoral preferences. In addition, the results can contribute to future lines of social interaction research on Facebook, for the understanding of the emotional dynamics that are generated in social networks by the content of propaganda in various social processes, such as elections or marketing campaigns to disseminate any product or service.

References

[1] "The Complete Guide to Facebook Analytics."

[2] Y. Tian, T. Galery, E. Molimpakis, and C. Sun, "Facebook Sentiment : Reactions and Emojis," pp. 11–16, 2017.

[3] Roberto Hernández Sampieri, *Metodología de la investigación*, vol. 6 edición. 2014.

[4] C. Integral and P. E. Federal, "Plan y Calendario Integral del Proceso Electoral Federal 2017-2018," 2018.

[5] A. Z. Jacobs, S. F. Way, J. Ugander, and A. Clauset, "Assembling thefacebook: Using heterogeneity to understand online social network assembly," 2015.

[6] W. They and S. It, "Contagious Content Marketing," pp. 1–35, 2014.

[7] C. Kim and S. U. Yang, "Like, comment, and share on Facebook: How each behavior differs from the other," *Public Relat. Rev.*, vol. 43, no. 2, pp. 441–449, 2017.

[8] P. K. Novak, "Sentiment of Emojis - EMOJI FACTS."

[9] P. Al and I. D. E. Junio, "M ÉXICO 2018 : E N LA RECTA FINAL ; LAS," vol. 2018, 2018.