

RESEARCH TOPICS OF THE COMPLEX SYSTEMS GROUP

SEPTEMBER 15, 2020

tCSG

PEOPLE

This document is intended for consumption by interested prospective BPhil, placement, Honours, Master and PhD students looking for projects in the general area of “Complex Systems”. This is a list of current, progressing or hoped-for research projects which we think are potential interesting.

Current members of the CSG - and therefore potential contacts, along with broadly stated research areas are as follows:

Ayham Zaitouny

Positioning and tracking; complex systems; nonlinear time series; dynamical systems; geological, ecological and biological systems

@ayham.zaitouny@uwa.edu.au

Brendan Florio

Industrial modelling, minerals processing, particle dynamics

@brendan.florio@uwa.edu.au

David M. Walker

Complex networks, Nonlinear time series, Dynamical systems, Modelling animal behaviour, granular media and other physical, ecological and biological systems

@david.walker@uwa.edu.au

Debora Correa

Nonlinear time series, Machine learning, Dynamical systems, Complex systems

@debora.correa@uwa.edu.au

Ismael Mola

Computational Fluid Dynamics

[ismael.mola@csiro.au](mailto:@ismael.mola@csiro.au), ismael.mola@uwa.edu.au

Leo Portes

Complex systems-driven data science, Differential network analysis, Synchronisation, whatever is necessary to find a needle in a haystack of data

@leonardo.portesdossantos@uwa.edu.au

Michael Small

Complex systems, Complex networks, Non-linear dynamics, Nonlinear time series, Complex data modeling

@michael.small@uwa.edu.au

Shannon Algar

Swarming, Swarm-Reservoir Computers

@shannon.algar@uwa.edu.au

Thomas Jüngling (T2)

Synchronisation, Reservoir computing, Complex systems, Dynamical systems

@thomas.jungling@uwa.edu.au

Thomas Lymburn (T3)

Reservoir computing, Synchronisation, Time series modelling

@thomas.lymburn@research.uwa.edu.au

Thomas Stemler (T1)

Complex systems and networks, Non-linear dynamics, Complex data modelling, paleo-climate analysis

@thomas.stemler@uwa.edu.au

PROJECTS

RANDOMISATION OF TIME SERIES VIA SHUFFLING

Debora and Jack developed an algorithm to shuffle symbolic sequences to obtain a random realisation of a particular Markov process ([Markov surrogate paper](#)). One could extend this to shuffle sequences of points in the time series generating the symbolic sequence in the first place. This would essentially provide a surrogate consistent with the hypothesis that the dynamical state information is encoded in the symbolic representation of the data. And therefore provide a way of testing whether ordinal network representatives (or any of the network approaches) actually are a good idea.

 *Small, Correa*

SWARMS AND RESERVOIR COMPUTING

This is the burgeoning T2+T3+Shannon collaboration. [The aims of this work are two-fold](#): use reservoirs to learn more about swarms; and use the intelligence of swarms to learn more about reservoirs. Both sides of this exploration address fundamental questions such as: how does nature evolve good swarms and how can we design one?; what is a good reservoir for computation?; and what is the information processing capacity of these systems? There is plenty of room for new angles in this too. Easy to define a distinct PhD project there. Here a swarm is a collection of interacting agents behaving as a flock. The warm acts as an alternative pattern generator for the machinery of the reservoir computer – or as an application for the predictive power of the reservoir.

 *T3, Algar*

RESERVOIR COMPUTING THEORY

Several possible projects for any level with the goal of explaining how information-processing performance emerges in driven dynamical systems. Big picture: Understanding of computation in natural systems, exploring physical limits of information processing, and the design of novel electronic/optical circuits and devices based on dynamical systems. One main branch of research is consistency theory (home-made characterisation of signal and noise in the reservoir, closely related to canonical component analysis, see [here](#)). Further questions involve analysis of the nonlinear features generated by the reservoir, and the role of the network structure of the reservoir for its function. We aim at a theory that integrates the entire scheme, from system parameters via the nonlinear response to the synthesis of target signals (supervised-learning).

 *T2*

INFORMATION THEORETIC MODELLING

There are techniques [developed by UWA](#) in the mid-to-late 90's for information theoretic constraints on model selection – the minimum description length stuff we talked about in MATH4021. This has recently been re-coded in Julia and, apart from providing a much faster implementation of 25 year old technologies, does raise some interesting questions. Specifically, one could apply the same model selection criteria to the output layer of a reservoir. Does a RC perform as well as RBF models? Features of the models could also be explored – basis functions could be represented as nodes, that are linked if they overlap, and then one has a neat abstract visualisation of the model structure. How does that help us understand the observed dynamics? How does one visualise such a model?

 *Small*

TOPOLOGICAL FEATURE EXTRACTION FROM NETWORKS

Stuff like persistent homology and topological data analysis are tool kits that provides a neat new way of quantifying the structural complexity of networks. It has not yet properly been applied to networks constructed from attractors, but it would be (maybe) [easy to do so](#). This would allow us to use network features to describe dynamics and – in particular bifurcation. We have methodologies to transform time series data into a network and for that network to represent the deterministic dynamical rule of the original system. That is, the network encodes the important dynamical features. However, some of these networks are large and it would be useful to be able to adapt the existing mechanism of topological data analysis (a fancy way of computing homology and

counting the holes in a point cloud) to better quantify these topological network features that map onto aspects of the underlying dynamics.

 Walker, Small

TIPPING POINTS

We are very interested, for a number of reasons that are mostly interconnected, **in when dynamical systems reach a state transition**. Simple techniques (std. dev. etc.) exist to quantify the loss of stability, but it is an open and unaddressed question. Can the network approaches or the description length modelling, or the RC ideas help? Is the **quadrant scan method sufficient**?

 Zaitouny, Small

NETWORK SURROGATES

This is an old itch that I haven't been able to scratch (it was a question in MATH4021). Essentially how do we randomise a given network to create random realisation of that network that are also consistent with specified hypotheses?

 Small

SWARM MODELLING

Based on previous work on the selfish herd that is a model to explain flocking and collective dynamics between agents (insects, birds, fish, etc.). Extending the selfish herd model to three dimensions and introduce predator-prey interaction as well as pinning to local attraction (shelter/food). Eventually, build, program and deploy an army of autonomous drones.

 Algar, Stemler

PALEO-CLIMATE TIME SERIES ANALYSIS

Paleo-Climate records, so called proxies, are often challenging to analyse. Beside of being noisy, they often have a non-uniform time sampling. Novel time series analysis methods are needed to deal with these complex data sets. These methods include but are not limited to machine learning methods, ordinal partitions, network based time series methods and transformation cost algorithms. Research is done in collaboration with the Potsdam Institute for Climate Impact Research, Germany and the NOODS group at K. Has University (Istanbul). HDR students will have to opportunity to visit either of these research groups.

 Stemler

PROPAGATION, CHASING, HIDING AND EXPLORING NETWORKS

Done lots of work on this over the years, probably a good time to look at it more **carefully**. What are the best ways for an agent to act to explore a network? To hide from a **hunter agent**? To spread disease? Etc.?

 Small

CONCEPT DRIFT DETECTION IN TIME SERIES USING COMPLEX NETWORKS

The representation of scalar time series as a complex network has proved to be useful to derive insights about the underlying process generating the data. The majority of existing techniques, however, do not scale well with the length of the time series, leading to very large networks. Our group has developed a strategy in which a compression algorithm is applied to the time series and its output is used to generate a complex network. The data compression algorithm exploits recurrences in the time series leading to more compact representations of the complex network coined a compression network. In this project, we want (i) generate compression networks in a windowing basis in order to detect concept drift in time series, that is, develop strategies that are capable of relating changes in compression network topologies to changes in the behaviour of the original time series, and

(ii) extend our compression network to multivariate time series. The successful application of this framework can be directly applied to many industrial scenarios, for instance, anomaly detection and predictive analysis.

 Walker, Small, Correa

INTEGRATING NETWORK ANALYTICS WITH MACHINE LEARNING

Researchers have been motivated to apply unsupervised techniques, another paradigm in ML—also known as *clustering*, to detect concept drifts in time series, that is, models that are able to detect changes in the time series behaviour. Clustering techniques do not require the existence of labels, and observations are grouped in such a way that observations in the same group are more similar than observations among groups. The usual approach is to build up models for consecutive data windows and use differences between the models to point out when changes occurred. In this project, we plan to integrate clustering techniques with the temporal features extracted from segments of time series and features extracted from temporal networks to track changes in the behaviour of multivariate time series.

 Walker, Small, Correa

All of this is largely methodological.

APPLICATIONS

The chief application domains that we are interested in – if this would help formulate your ideas better – are:

YOUNG LIVES MATTER

a suicide prediction project that I help run. There is a separate project looking at indigenous mental health that we are also involved with. Both have data, mostly really nasty data, and need new methods.

 *Small*

THE MAINTENANCE INDUSTRIAL TRANSFORMATION TRAINING CENTRE

the data science of maintenance industry project that Michael, Ed, Adriano, Debora and Ayham are connected with. The tagline is “transforming maintenance through data science”. Our interest here is in applying machine learning, statistics and dynamical systems/complex systems theory to predict failure in maintenance equipment. Funded by our industry partners BHP, Roy Hill and Alcoa.

 *Correa, Zaitouny, Small*

A NEW OFFSHORE DATA ENGINEERING TRAINING CENTRE

just been announced, similar to the preceding thing, but in the water. Funded by a different bunch of Oil and Gas companies.

 *Small*

ANIMAL WELFARE VIA TIME SERIES AND (POSSIBLY) NETWORK MEASURES

Dave, Debora and Me are investigators (along with colleagues from animal/biology on a project to evaluate stress in animals from various sensors).

 *Algar, Walker, Correa*

TRAFFIC

A project that T1 is leading for us looking at optimal traffic and road design to improve traffic flow. Methods include Shadowing filters, machine learning for big data, control theory, sensitivity analysis and information based modelling.

 *Stemler*

BIOINFORMATICS AND GENE EXPRESSION

the data that Leo is looking at with some people from cancer medicine to design improved “personalised” medicine.

 *Portes, Zaitouny*

CANCER COUNCIL

Use mathematics and complex systems to investigate dynamics changes over time on sequencing data that can reveal useful information about the nature of the interactions between the immune system and the cancer. Basically, we aim at understanding the reason some individuals respond to cancer immunotherapy while others do not. Similar to the Leo gene expression stuff, but not identical.

 *Correa, Zaitouny*

NATIONAL GEOGRAPHIC

A project that Debora is working on to investigate the impact of forestation/deforestation in temperature extremes in Oceania and Indonesia's regions. We will be using deep learning and other ML models.

 *Correa, Small*

DISEASE MODELLING

not yet funded, but there are two potential projects here, one with colleagues in computer science and medicine at UWA, and another with Hong Kong. The primary interest is in **network models of disease propagation**.

 *Small*

ML AND OPTIMISATION METHODS TO ESTIMATE DIET COMPOSITIONS IN UNGULATES.

Ungulates are the link between producers and consumers in the trophic chain with the interface between domestic and wild species becoming a problem of growing interest. Traditional theory on food resource partitioning is based on large differences of body size between species, however, little is known on the structure and function of inter- and intra-species interactions. Despite the advantage of GPS technology to monitor ungulate interaction the logistics required (e.g. cost, limited sample size) pose challenges. There is potential to use nutritional indices to assess interactions by developing proxies of them to unravel species interactions. Specifically, formalizing methods based on natural plant markers (n-alkanes from faecal matter) to estimate diet composition and overlap without the need of knowing exactly the plants in the diet. By inferring such compositions using ML methods we hope to be able to develop a non-invasive tool to compare diets of species and from knowing the ecological footprint of the landscape unravel interactions between and within species.

 *Walker, Correa*

THE COVID-19 CARE STUDY OF MENTAL WELLBEING DURING PHYSICAL DISTANCING: NATURAL LANGUAGE PROCESSING APPROACHES TO PREDICTING SOCIAL BEHAVIOUR AND MENTAL HEALTH

The COVID-19 CARE Study is a 14-day online survey study led by Forrest Fellow Dr Julie Ji and Dr Debora Correa from the School of Psychological Science, UWA, in collaboration with the Forrest Research Foundation and Minderoo Foundation. The study has collected data on key demographic, cognitive, and emotional variables from a large community sample of Australians to investigate the extent to which physical isolation predicts social isolation and loneliness, in whom, and why. Natural language data from N = 1000+ participants include 1) brief open-ended responses on top COVID-19 related concerns and silverlinings; and 2) paragraph-sized descriptive answers from a mental problem-solving exercise. This research project offers the unique opportunity to use NLP approaches to understand how characteristics of thought content predict social behaviour, loneliness, and mental health during COVID-19.

 *Correa*
