A Bayesian Penalized Hidden Markov Model for Ant Interactions

Meridith Bartley[†], Ephraim Hanks[†] and David Hughes

Abstract. Interactions between social animals provide insights into the exchange and flow of nutrients, disease, and social contacts. We consider a chamber level analysis of trophallaxis interactions between carpenter ants (*Camponotus pennsylvanicus*) over 4 hours of second-by-second observations. The data show clear switches between fast and slow modes of trophallaxis. However, fitting a standard hidden Markov model (HMM) results in an estimated hidden state process that is overfit to this high resolution data, as the state process fluctuates an order of magnitude more quickly than is biologically reasonable. We propose a novel approach for penalized estimation of HMMs through a Bayesian ridge prior on the state transition rates while also incorporating biologically motivated covariates. This penalty induces smoothing, limiting the rate of state switching that combines with appropriate covariates within the colony to ensure more biologically feasible results. We develop a Markov chain Monte Carlo algorithm to perform Bayesian inference based on discretized observations of the contact network.

Keywords: Bayesian estimation, hidden Markov model, penalization, ant trophallaxis.

1 Introduction

Penalized estimation has long been a valuable tool in the development of models that provide biologically reasonable and interpretable results. Ridge and LASSO penalties provide a means for variable selection and regulation in a linear model framework. However, penalized estimation in the context of stochastic processes has yet to be fully explored. When exploring second by second ant behavior data we observed evidence of switching between periods of high and low rates of feeding interactions. When a standard Poisson hidden Markov model was applied to these data the resulting latent state predictions were overfit, or switching much faster than biologically reasonable. That is, we would not expect a chamber- or colony-level behavior to switch on the order of seconds. We propose an approach to penalize stochastic processes within a hidden Markov model framework and apply this approach to high resolution behavioral interaction data. Our work provides a way to combat overfitting common in high resolution data through Ridge or LASSO-like penalizing priors within a Bayesian context.

1.1 Ant System

Ants provide an ideal system for which to study social organisms in a controlled environment. Ants can be used as a proxy for learning about how humans in a social

---

[†]Department of Statistics, Penn State University bartley@psu.edu hanks@psu.edu
[†]Department of Entomology, Penn State University hughes@psu.edu

community deal with spread of disease, space utilization, and environmental variation (Fewell (2013)). While we cannot conduct ethical controlled experiments on human populations, ant colonies are easy to maintain and manipulate in a laboratory setting. Researchers can control the space (size and organization), infections, nutrients, etc all while continuously monitoring behavioral interactions within the colony. By understanding how ants behave in these systems we may draw conclusions about our own interactions and abilities to mitigate spread of diseases.

Ants engage in an oral exchange of nutrients called trophallaxis. Much like a hand shake or embrace in our own communities, these interactions in ants are a primary opportunity for disease transmission (Naug and Camazine (2002)) and for sharing nutrients through the colony. On both the colony and functional group level, ants have been shown to organize themselves, both spatially and temporally, in a way that defends against the spread of disease (Quevillon et al. (2015)). It is imperative to be able to model interactions first under a healthy system to provide a basis for comparison when subsequent work manipulates the colonies through experimental infections. The Hughes Lab at Penn State maintains several wild-collected colonies of the black carpenter ant, *Camponotus pennsylvanicus*, continuously monitored inside specially designed wooden chambers under complete darkness. Videos were examined frame by frame to identify and record each feeding event, the individual ants involved, and the start and end timing. In this study, we use four hours of these continuously monitored feeding interactions for one colony, and consider only interactions observed in the nest chamber in which the queen ant resides. The data points within Figure 1 shows the cumulative number of trophallaxis interactions observed in the four hours (14,400 seconds) of observations within the single chamber.

An initial examination of the feeding interactions over time (see Figure 1) suggests there may be periodic pulses in the rates of events. These pulses may reflect some latent underlying chamber-level behavioral states that are driving the different trophallaxis rates over time. For example, foraging ants re-entering the nest (as allowed in this experimental setup) may spark a string of trophallaxis events as they seek to distribute nutrients through the colony. Previous research has also shown that ant colonies have been found to exhibit collective activity cycles (Richardson et al. (2017)). However, although we may observe evidence of these cycles, the mechanisms or causes of switching remain uncertain. The combination of high-resolution observations of each ant feeding exchange with these proposed unobserved chamber-level behavioral states lends itself nicely to exploration via a hidden Markov model. Hidden Markov models are often used to model ecological systems because they allow for behaviors to be correlated over time in a way that accounts for shifts in an underlying state process. HMMs have been used recently to study animal movement behavior (Langrock et al. (2012); McKellar et al. (2015); Patterson et al. (2017); Towner et al. (2016); van de Kerk et al. (2015)), general animal behavior (DeRuiter et al. (2016); Langrock et al. (2014); Schliehe-Diecks et al. (2012)), as well as other applications within population ecology (Borchers et al. (2013); Gimenez et al. (2014); Johnson et al. (2016); Leos-Barajas et al. (2017)).

When fitting a two-state Poisson HMM to the ant feeding interaction observations (described below in Section **??**), the estimated state switching rates were much too
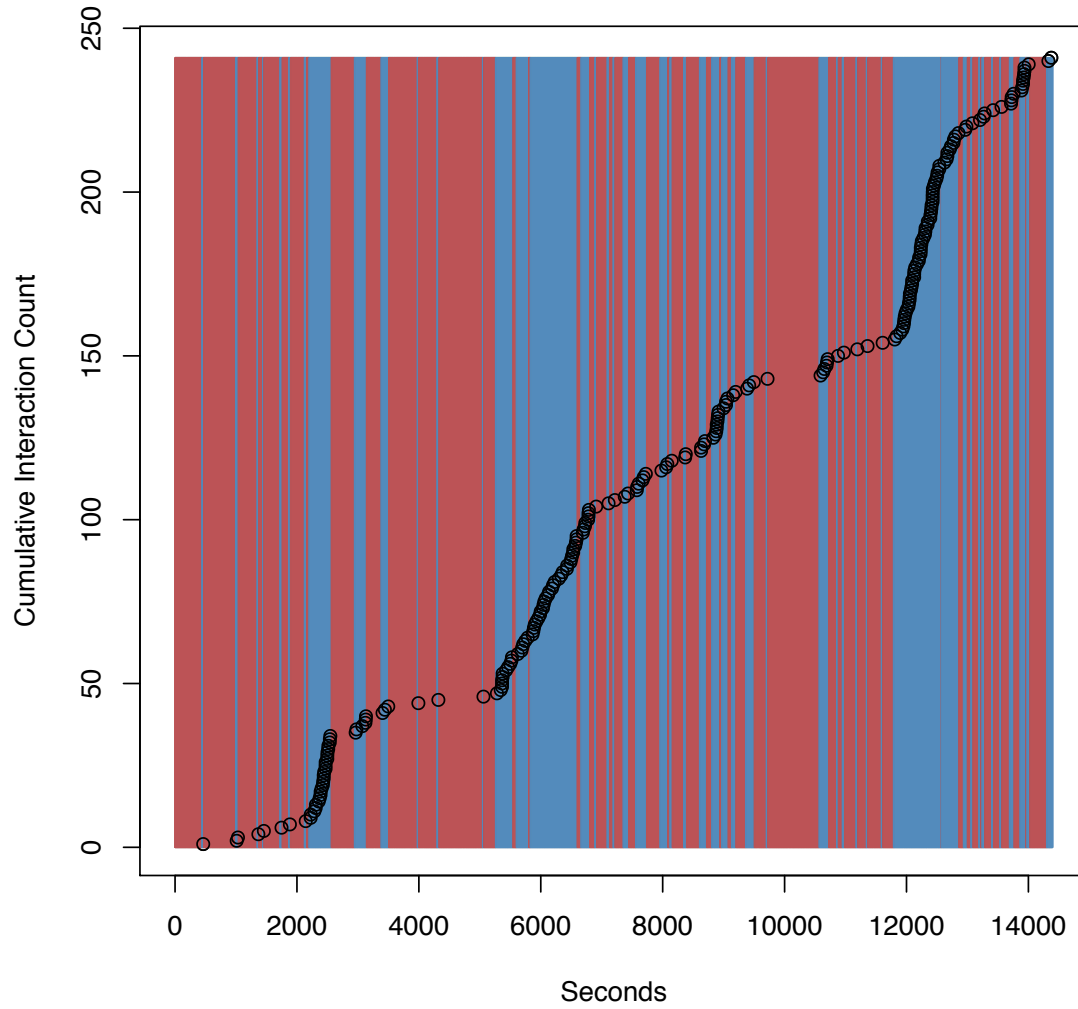
Figure 1: While periods of relatively low and high rates of feeding interactions are clear, the stochastic process of chamber-level feeding states is switching between the two much faster than biologically feasible. Red background denotes low state while blue background denotes high state of feeding exchanges.

fast to be biologically reasonable (Figure 1) (switching approximately 0.8 times per minute). In this Figure, the background colors denote the estimated chamber level behavior state. Red background denotes a predicted low state while a blue background denotes the high state of feeding exchanges. In fact, the best estimated latent chamber level state process often switches between states immediately before and after individual interaction events resulting in fast estimated switching rates. This motivates the need to develop penalized approaches to fitting stochastic processes, like HMMs, to combat the apparent overfitting of the data. High resolution time series are more and more common, so this problem will be important in many future scientific applications.

What is typically done in the literature to resolve an issue of short "state-dwell times" (i.e. the number of consecutive time points that the Markov chain spends in a given state) is to move to a hidden semi-Markov model (HsMM), which can be expressed in discrete time. While a standard HMM dwell times follow a geometric distribution, Hidden semi-Markov models relax this condition; the dwell time in each state can follow any discrete distribution on the natural numbers (CITE YU2009, LANGROCK AND ZUCCINI 2010). However, this approach does not easily allow for the incorporation of covariates. While Langrock and Zucchini propose restructuring HsMM within a specialized HMM design to take advantage of all of the HMM framework. (WHY OURS OVER THIS APPROACH?)

In this paper, we propose a novel penalized hidden Markov model that penalizes rapid state switching through Bayesian ridge- or LASSO-like priors on continuous-time Markov chain transition rates. We show that this smooths the rate of estimated behavioral state switching and results in better 1-step ahead prediction than the non-penalized HMM. We demonstrate the utility of this penalized approach to stochastic process modeling on the ant trophallaxis data. We then extend this approach to model the effect of biological covariates, such as the arrival of forager ants into the colony or chamber, on transition rates between HMM states in order to better understand the cause of these behavioral switches.

References

Borchers, D. L., Zucchini, W., Heide-Jørgensen, M. P., Cañadas, A., and Langrock, R. (2013). "Using Hidden Markov Models to Deal with Availability Bias on Line Transect Surveys." *Biometrics*, 69(3): 703–713. 2

DeRuiter, S. L., Langrock, R., Skirbutas, T., Goldbogen, J. A., Chalambokidis, J., Friedlaender, A. S., and Southall, B. L. (2016). "A multivariate mixed hidden Markov model to analyze blue whale diving behaviour during controlled sound exposures." 1–26.
URL http://arxiv.org/abs/1602.06570 2

Fewell, J. H. (2013). "References and Notes." *The Hunt for the Parathyroids*, 301(September): 132–143.
URL http://doi.wiley.com/10.1002/9781118343371.refs 2

Gimenez, O., Blanc, L., Besnard, A., Pradel, R., Doherty, P. F., Marboutin, E., and

Choquet, R. (2014). "Fitting occupancy models with E-SURGE: Hidden Markov modelling of presence-absence data." *Methods in Ecology and Evolution*, 5(6): 592–597. 2

Johnson, D. S., Laake, J. L., Melin, S. R., and DeLong, R. L. (2016). "Multivariate state hidden Markov models for mark-recapture data." *Statistical Science*, 31(1958): 233–244. 2

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., and Morales, J. M. (2012). "Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions." *Ecology*, 93(11): 2336–42.
URL http://www.ncbi.nlm.nih.gov/pubmed/23236905 2

Langrock, R., Marques, T. A., Baird, R. W., and Thomas, L. (2014). "Modeling the Diving Behavior of Whales: A Latent-Variable Approach with Feedback and Semi-Markovian Components." *Journal of Agricultural, Biological, and Environmental Statistics*, 19(1): 82–100. 2

Leos-Barajas, V., Photopoulou, T., Langrock, R., Patterson, T. A., Watanabe, Y. Y., Murgatroyd, M., and Papastamatiou, Y. P. (2017). "Analysis of animal accelerometer data using hidden Markov models." *Methods in Ecology and Evolution*, 8(2): 161–173. 2

McKellar, A. E., Langrock, R., Walters, J. R., and Kesler, D. C. (2015). "Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird." *Behavioral Ecology*, 26(1). 2

Naug, D. and Camazine, S. (2002). "The role of colony organization on pathogen transmission in social insects." *Journal of Theoretical Biology*, 215(4): 427–439. 2

Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., and King, R. (2017). "Statistical modelling of individual animal movement: an overview of key methods and a discussion of practical challenges." *AStA Advances in Statistical Analysis*, 101(4): 399–438. 2

Quevillon, L. E., Hanks, E. M., Bansal, S., and Hughes, D. P. (2015). "Social, spatial, and temporal organization in a complex insect society." *Scientific Reports*, 5: 13393.
URL http://www.nature.com/doifinder/10.1038/srep13393 2

Richardson, T. O., Liechti, J. I., Stroeymeyt, N., Bonhoeffer, S., and Keller, L. (2017). "Short-term activity cycles impede information transmission in ant colonies." *PLoS Computational Biology*, 13(5): 1–17. 2

Schliehe-Diecks, S., Kappeler, P. M., and Langrock, R. (2012). "On the application of mixed hidden Markov models to multiple behavioural time series." *Interface Focus*, 2(2): 180–189.
URL http://rsfs.royalsocietypublishing.org/cgi/doi/10.1098/rsfs.2011.0077 2

Towner, A. V., Leos-Barajas, V., Langrock, R., Schick, R. S., Smale, M. J., Kaschke, T., Jewell, O. J., and Papastamatiou, Y. P. (2016). "Sex-specific and individual pref-

erences for hunting strategies in white sharks." *Functional Ecology*, 30(8): 1397–1407.
2

van de Kerk, M., Onorato, D. P., Criffield, M. A., Bolker, B. M., Augustine, B. C., Mckinley, S. A., and Oli, M. K. (2015). "Hidden semi-Markov models reveal multiphasic movement of the endangered Florida panther." *Journal of Animal Ecology*, 84(2): 576–585. 2