DOI: http://dx.doi.org/10.7551/978-0-262-32621-6-ch054

Inferring Social Structure of Animal Groups From Tracking Data

Brian Hrolenok[1], Hanuma Teja Maddali[1], Michael Novitzky[1], Tucker Balch[1] and Kim Wallen[2]

[1]Georgia Institute of Technology, Atlanta, GA 30332
[2]Department of Psychology and Yerkes National Primate Research Center, Emory University, Atlanta, GA 30322

Abstract

Inferring the social structures of animal groups from their observed behavior is a non-trivial task usually handled by direct observation. Recent advances in sensing and tracking technology have enabled the collection of dense spatial data over long periods of time automatically. The qualitative differences between sparse hand-coded data and dense tracking data necessitate a new approach to inferring the social structure of the observed animals. We present a framework for using agent-based simulations to guide our approach to inferring social structure from tracking data collected from a small group of rhesus macaques over a period of three months. As part of this framework, we describe a version of the DOMWORLD model of dominance interactions in rhesus macaques that has been modified to include association preference, and adapted to more closely match the environment where the monkeys were housed. An exploration of simulation results reveals important characteristics of the tracking data. The inferred social structures of the tracked monkeys are also presented.

Introduction

Biologists and psychologists studying the social structures and dynamics of animals have relied on observation by trained researchers for the collection and coding of data, as until recently automated tracking systems have not been able to provide the accuracy required to recognize events of importance. With the advent of new tracking methods and subsequent improvements in tracking accuracy, it is now possible to record accurate, high-frequency spatial information over long periods of time. This qualitatively different kind of data requires a new approach to analysis.

As the sheer volume of data prohibits manual analysis, automated methods are necessary both for identifying key events and inferring relevant characteristics from identified events. In the rest of this paper, we examine how such automated methods can be applied in the specific case of inferring the social structure of a group of six rhesus macaque monkeys given tracking data of their movements over a period of three months at a rate of about 30Hz. In the next section we will review some related literature on social structure and agent-based modeling. Following that we will highlight the specific aspects of social structure we are interested in recovering, and related behaviors. Next, we lay out our motivation for using agent-based modeling in this work. After that we cover the details of our approach to modeling, and inferring social structure. After that we present some results using simulated and real data. Finally, we provide some high level analysis, conclusions, and directions for future work.

Motivation and Related Work

Social structure in primate groups plays an important role in the health, behavior, and development of group members. Wallen (1996) has shown that social structure plays an important role in the development of behavioral sex differences, while Stephens and Wallen (2013) describe how social status can effect the actual physiological development of young monkeys. Sapolsky (2005) reviews how social status can effect a wide range of health issues, both direct (such as access to resources), and indirect (stress related diseases). Being able to automatically infer the social structure of a group of animals then has wide ranging implications from maintaining the health and safety of laboratory animals, to determining the changes in social structure throughout the course of an experimental protocol.

In order to guide our development of automatic algorithms for inferring social structures from dense tracking data, we take an agent-based modeling approach to creating simulations of animal behavior in order to prototype and refine our methods. The work presented by Yang et al. (2012) has a similar goal, and provides a principled framework for using agent-based models to further the ethological study of foraging behaviors, specifically the foraging behavior of *Aphaenogaster cockerelli*. Hrolenok and Balch (2013a) presented work on learning these agent-based models of ant foraging directly from data using techniques from machine learning, and later (2013b) fish schooling, although there the purpose was the automated learning of the behavior model itself, while in this work we are interested in developing inference techniques using a known model. Our development of this known model is heavily influenced by DOMWORLD, introduced by Hemelrijk (2000). Hemelrijk

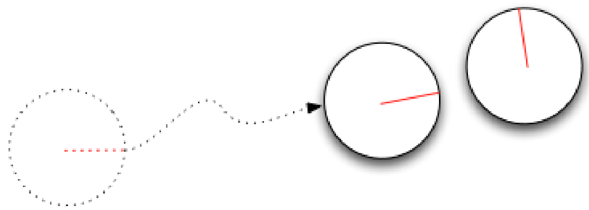


Figure 1: An approaching behavior that indicates an association preference. The strength of the indicated association preference is determined by the frequency and length of periods of close proximity.

presents an agent-based model of dominance in primates that emerges spatial patterns typically found among certain types of rhesus macaque troops.

Social Structures in Rhesus Macaques

One of the most intuitive measures of social structure in primates is association preference, which indicates which members of the group each individual prefers to spend time in close proximity to. A graph constructed from association preferences can illuminate subgroups, key individuals which connect otherwise disconnected groups (also known as cut vertices, or articulation points), as well as overall measures of group structure such as connectedness. The observable behavior where two or more monkeys spend time within relatively close proximity to one another indicates association preference. Figure 1 illustrates a grouping behavior that indicates association.

Another important measure of social structure is the dominance hierarchy. Dominance plays important roles both in interactions between individuals and group dynamics, and changes in dominance can indicate significant events of interest to the primate researcher. Observed displacement and withdrawing behaviors such as chasing and fleeing indicate a dominance relationship. Figure 2 illustrates a withdrawal behavior that indicates a dominance relationship.

While some association, displacement, or withdrawing behaviors can only be identified visually, a large number of them can be detected directly from spatial data, as described in later sections. In order to obtain this data, we utilized a 3D position tracking system to track the positions of 6 monkeys in a 3m x 3m enclosure over a period of three months. Details of the tracking system are described in Huang et al. (2012).

The Importance of Agent Based Models

Agent-based modeling and simulation of animals solve two major problems in the experimental study of animal behavior. First, the data collection cost associated with studies

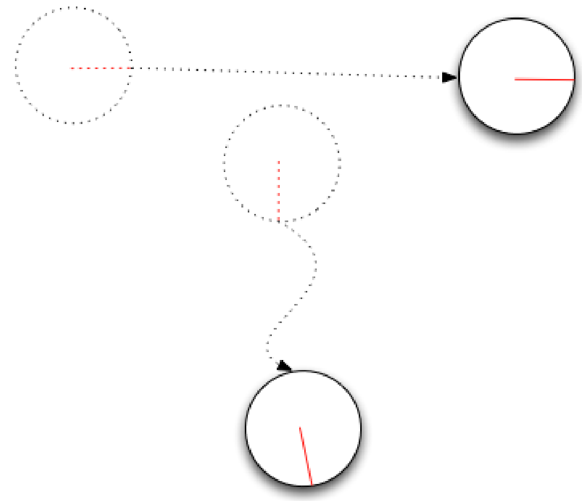


Figure 2: A withdrawal behavior that indicates an dominance relationship. The strength of the indicated dominance relationship is determined by the relative frequency with which each individual withdraws from the other.

done in simulation using high fidelity models is essentially zero, at least compared to the cost of running experiments and collecting data on real animal subjects. Using ABMs allows the researcher to run simulated experiments to increase the confidence of statistical analysis which might otherwise be less conclusive.

Second, when inferring model parameters directly from data, one is faced with the task of validation without access to any "ground truth". Performing the same inference methods on simulated data can provide crucial insight into how those techniques may perform on data from live animals. Both success and failure can be valuable clues into the capabilities and limitations of the inference methods.

In this work, we focus on using ABMs as a validation tool. In the next section we introduce an agent-based model of social interactions between monkeys based on a well studied simulation with slight modifications relevant to the specific social measures mentioned previously. Using this ABM, we can measure quantitatively the effectiveness of our methods for recovering social structure.

Methodology

To validate our method for inferring social structure, we created an agent-based model that incorporates the important behaviors mentioned previously, parameterized in a way that allowed us to compare the recovered social structures with the "ground truth" of the simulation.

Agent Based Behavior Model

Our simulation model, which we call SMALLDOMWORLD is a modification of the earlier DOMWORLD of Hemelrijk

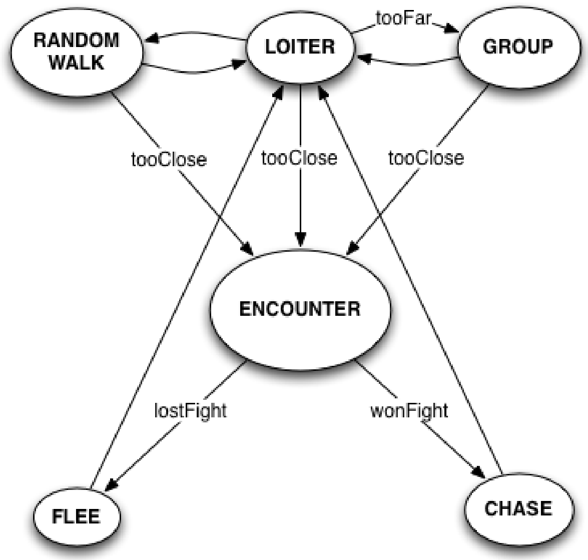


Figure 3: The SMALLDOMWORLD model.

(2000). The behavior of the individuals is guided by three components: a grouping component that draws individuals together, a dominance component where individuals confront each other and the winner chases the loser, and a random component where individuals wander about their environment at low speed.

In order to match the environmental conditions of the animals being studied, we modified the DOMWORLD model presented in Hemelrijk (2000) in a number of ways. In the troop we studied, the dominance relationships were already stable and established, where as in DOMWORLD, dominance relationships are recalculated after every encounter. In DOMWORLD, as in our model, dominance encounters only occur when an individual approaches another within some distance threshold representing an intrusion into personal space. In our model, the probability of an intrusion on personal space resulting in a dominance encounter is given by the parameter σ where $\sigma = 1.0$ indicates a completely stable dominance structure with no confrontations, and $\sigma = 0.0$ ensures that any intrusion results in a confrontation. We use the same dominance confrontation mechanism as DOMWORLD: each individual is given a dominance weight, and the difference in weights probabilistically determines the winner of any encounter (see Hemelrijk (2000) for details).

We also introduced some selectivity into the grouping behavior. Grouping in DOMWORLD represented a desire by all individuals in the group to remain within some proximity of other group members, and so when an individual found itself far away from the center of the group, it selected another visible member of the group uniformly randomly to head towards. In our model, each individual has a list of association preferences $\pi = \langle \pi_1, \pi_2, \ldots, \pi_n \rangle$ which can be thought of as the distribution over individuals selected for grouping. This leads to patterns of association which are non-uniform and give rise to the social structures described earlier. The list of association preferences can be combined into a single association preference matrix P with each row corresponding to a single individual's association preferences.

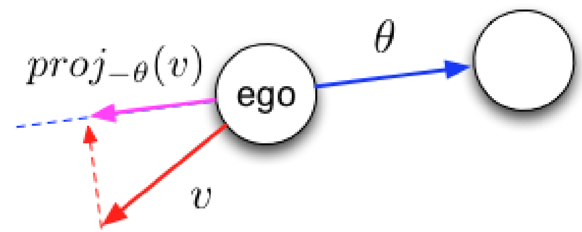


Figure 4: Detection of fleeing events.

While these two modifications represent the most important changes between our model and DOMWORLD, we also made several changes to accommodate differences in the modeled environment and simulation engine. DOMWORLD's environment is long-range, discrete, and unbounded (toroidal in implementation), whereas our environment is quite small, continuous, and interactions with the boundary of the enclosure are common (which necessitated some kind of obstacle avoidance behaviors). Figure 3 gives a graphical representation of the behavior model of individuals in SMALLDOMWORLD.

Heuristic Behavior Recognition

In this section we present two heuristic methods for identifying association and dominance behaviors, and how they can be used to infer the social structures of a group of monkeys.

Time spent within proximity is a straightforward way to detect behavior which indicates association preference among group members. By counting the frequency and length of events where the *ego* — by which we mean the individual whose preferences we are trying to determine — comes within a threshold distance of another individual, and remains there at low to zero velocity for at least some minimum period of time, we can infer which individuals the ego prefers to spend time with. If we denote by E_{ij} the time monkeys i and j spend near each other, then we can fill out the entries of the association preference matrix P as:

$$P_{ij} = \frac{E_{ij}}{\sum_k E_{ik}} \tag{1}$$

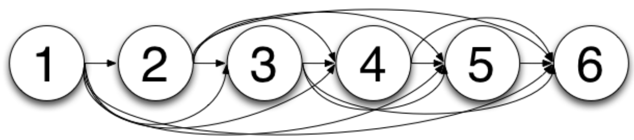


Figure 6: Dominance hierarchy.

While detecting all types of withdrawal events may be difficult, we can capture a certain subset fairly easily. One type of withdrawal involves the ego rapid moving directly away from the target, which we will call *fleeing* events. We can detect fleeing events by counting the length and number of events where the magnitude of the ego's velocity (v) projected onto the bearing (θ) between the ego and target is larger than a threshold (f), which is shown graphically in Figure 4. For each pair of individuals we can compute a dominance measure d as

$$d_{AB} = \left|\{e_{AB} \mid proj_{-\theta(A,B)}(v_B) > f\}\right| \quad (2)$$

If individual A flees from individual B more frequently than the opposite ($d_{BA} > d_{AB}$), we can infer that A is subordinate to B. In practice, picking f to correspond to roughly 30 degrees on either side of moving directly away from the target worked well.

Experiments

We performed four experiments to test our approach, three using data collected from our simulation model SMALLDOMWORLD with different parameterizations of the social structure, and one using the three months of tracking data we collected from a small group of animals. Our purpose in performing the simulation experiments was to measure how accurately we would be able to recover social structure, and to characterize scenarios where our method might not be able to recover social structure. In each of the simulation experiments we ran ten experiments under the same parameterization but with different initial configurations and random seeds.

In the first experiment, we simulated a group of monkeys which had a social structure with two disconnected subgroups, as shown in Figure 5. The parameterization which realized this structure had each individual's association preference set to 1.0 for other members of its subgroup, and 0.0 otherwise. This way there should be no deliberate preference to spend time in proximity of non-subgroup members. The dominance relationships for this and the two following simulation experiments was a direct linear relationship with rank corresponding to ID, as illustrated in Figure 6. The parameterization that realized this model set the dominance weight for the least dominant individual to 1.0, with each individual higher in the chain having twice the dominance

Table 1: Simulation parameters common for all experiments.

Personal distance	0.25m
Near distance	0.8m
Fleeing speed	2.0m/s
Chasing speed	1.0m/s
Grouping speed	0.25m/s
Wander speed	0.12m/s

Table 2: Frobenious error of recovered association preference as compared to a randomly generated symmetric, normalized matrix with zero diagonal. Averaged over 10 runs.

Recovered AP	Avg. error (std.)	random AP
disconnected	0.1744 (0.0014)	0.2408 (0.0326)
neutral hinge	0.1002 (0.0015)	0.1797 (0.0350)
preferred hinge	0.1388 (0.0004)	0.1869 (0.0158)

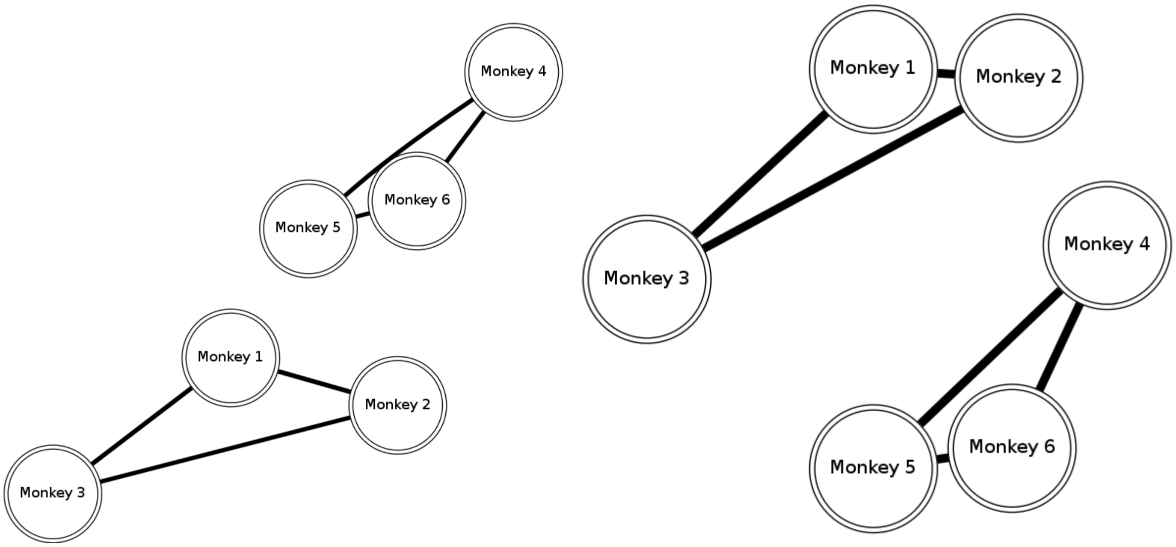
weight as the next lowest, or ($D = 2^{N-i}$). The same dominance weights were used in each experiment. Hierarchy stability was set fairly high ($\sigma = 0.8$), so that dominance interactions were not frequent, but still frequent enough to reliably detect the dominance hierarchy. Other simulation parameters are listed in Table 1 and were estimated from tracking data of live monkeys where appropriate, or taken from the literature when available.

We were successfully able to consistently recover both the dominance relationships and association preferences in this experiment. In order to get a sense of our accuracy, we compared the recovered association preference matrix with the ground truth parameterization and with a randomly generated matrix which was restricted to the same form (row-normalized, zero-diagonal, symmetric). The results of this comparison are given in the first row of Table 2, which shows that our recovered parameters are significantly closer to the ground truth than random. In order to recover the graph structure shown in Figure 5, we chose a threshold τ, such that edges larger than τ are included in the graph while those smaller are not. In order to characterize our ability to pick τ reliably, we examined the distribution of values in the association preference matrix P, shown in Figure 7. The distribution is clearly divided into two modes, which indicates that by picking a a threshold between the two modes, our recovered graph will be stable to noise in the estimation of association preferences. In our testing, picking

$$\tau = \frac{1}{n^2} \sum_{i,j} P_{ij} \quad (3)$$

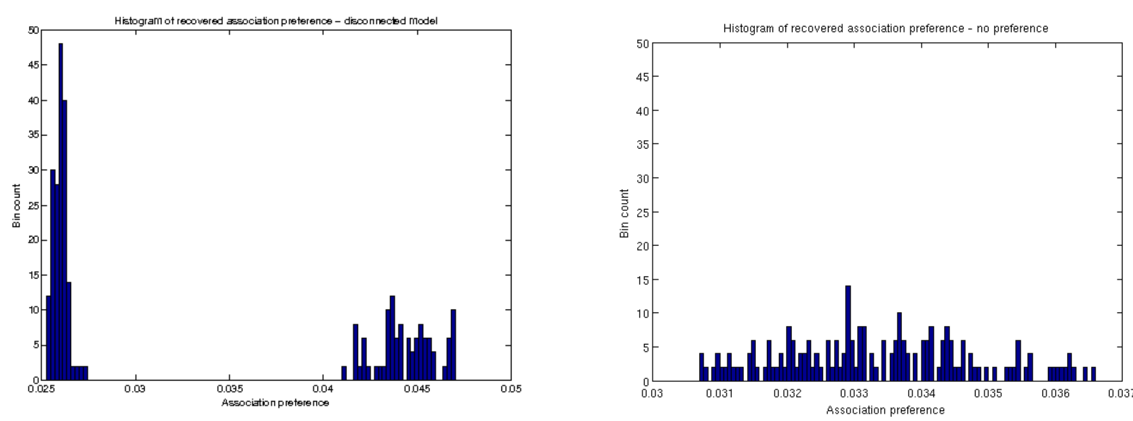
where n is the number of agents, worked reliably.

In the second experiment, we modified the association preferences so that one individual, which we will call the



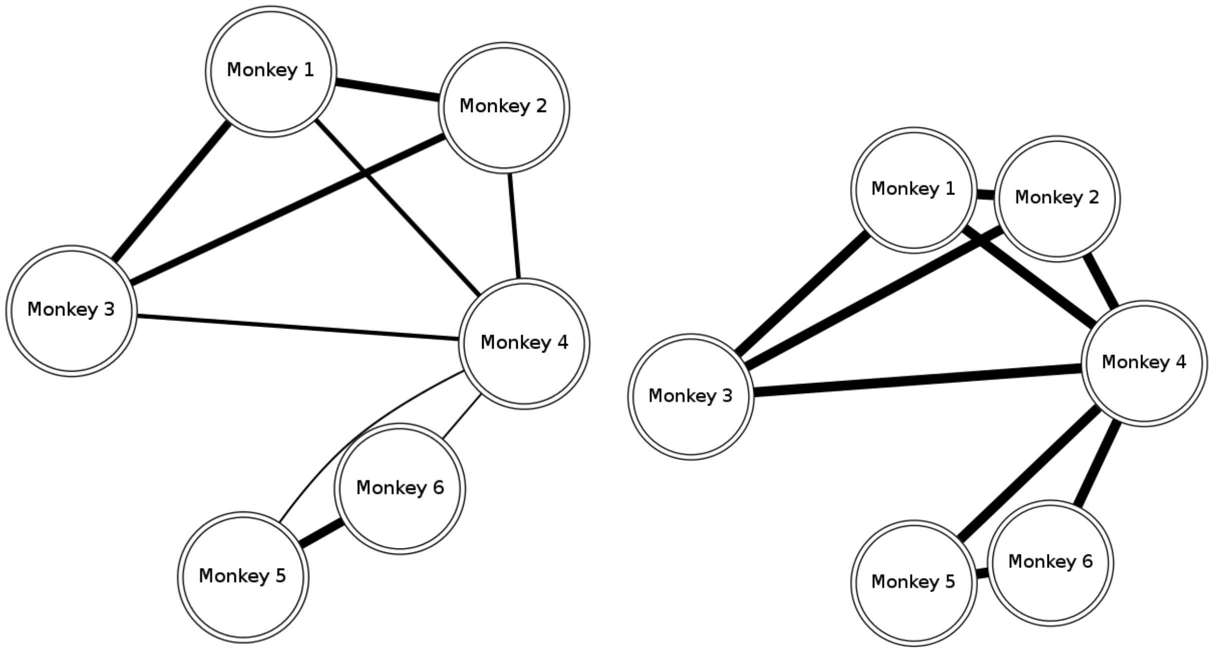
(a) Recovered association preference. (b) Ground truth association preference.

Figure 5: Association preferences for the disconnected scenario. The recovered graph (5a) closely matches the actual association preferences used in the simulation (5b). Line thickness corresponds to strength of association preference. Association preferences that fall below the threshold τ (from Equation 3) are not shown.



(a) (b)

Figure 7: Histogram of association preference values recovered (7a) from the disconnected scenario, and (7b) a simulation with no association preferences. In the second simulation, agents followed the same behavior model, except when choosing to group where they chose among neighbors without preference. Notice the clear separation into two modes of the recovered association preferences as compared to the noisy unimodal distribution from the simulation with no preferences.



(a) Recovered association preference.   (b) Ground truth association preference.

Figure 8: Association preferences with hinge node. The recovered graph (8a) closely matches the actual association preferences used in the simulation (8b). Line thickness corresponds to strength of association preference. Association preferences that fall below the threshold τ are not shown.

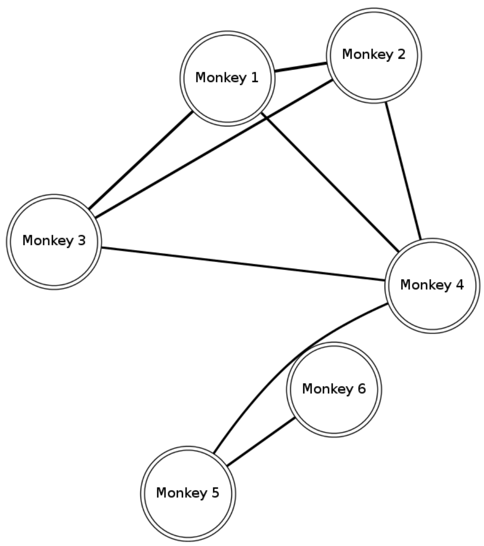
hinge, became an articulation point linking the two subgroups, as shown in Figure 8. To do this, we modified the association parameters such that the hinge individual preferred everyone equally, but no one had a preference for them. In terms of our parameterization, we set the hinge individual's row $P_{hj} = 1.0, \forall j$, and its column $P_{ih} = 0.0, \forall i$. Results are shown in the second row of Figure 2. Again, the recovered association preference is significantly closer to the ground truth than a random association preference, and the dominance hierarchy was recovered without error.

In the third experiment, we repeated the previous experiment, but also allowed the other individuals to preferentially group with the hinge individual by setting $P_{ih} = 1.0, \forall i$. By doing this we highlight a potential scenario where our method may not be able to recover the social structure accurately, specifically non-transitive association preferences. Note that our metric for association preference makes no distinction between individuals which are within proximity because they chose to be, and those that just happen to be nearby. For example, if individuals A and B do not have any preference for association, but each has a high preference for associating with a third party C, then *regardless* of C's preferences, A and B will spend a high proportion of time in proximity of one another. Figure 9 as well as the third row of Figure 2 illustrate how recovery performance is degraded in this type of non-transitive scenario.

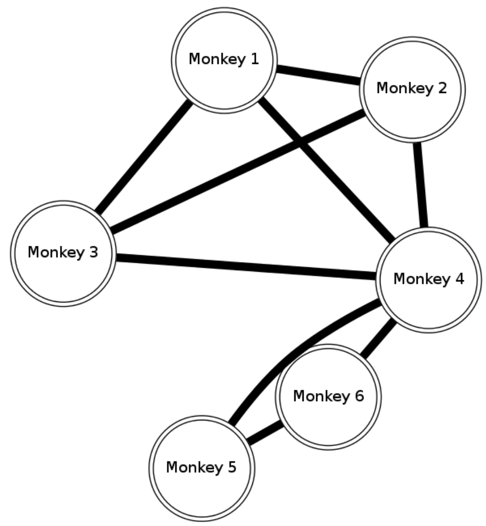
Finally, for our fourth experiment we applied our methods to tracking data of live animals. Figures 10 and 11 show the recovered association preferences and dominance hierarchy for the entire period the monkeys were tracked. We picked τ using the same approach described above, although from examining the distribution of association preferences shown in Figure 12, we know that this choice will be less stable with respect to noise. That is, it is more likely that some edges will be included or excluded from the graph due to small changes in association preference. The dominance relationship is a linear hierarchy $(4, 3, 5, 2, 6, 1)$ with individual 4 being the most dominant, and individual 1 being the least dominant. This agrees with the recovered association preference, where individuals are shown as preferring to associate with other individuals at similar ranks in the hierarchy.

Conclusion

We have described a framework for using agent-based models to explore the characteristics of automated techniques for analyzing dense spatial data. In the specific case of inferring the social structure of rhesus macaques from tracking data, we have illustrated under what conditions simple heuristic analysis can provide accurate reconstructions of social structure, and provided some insight into the stability and reliability of our approach using an extension of the DOMWORLD agent-based model to guide our analysis.



(a) Recovered association preference.



(b) Ground truth association preference.

Figure 9: Association preferences for the hinge node scenario using non-transitive preferences. The recovered graph (9a) is missing a link between monkey 6 and monkey 4 in the actual association preference graph (9b). Notice also that the magnitude of the preferences — shown by the thickness of the edges — is much closer to the threshold value τ. Picking smaller τ results in additional edges that are not present in the simulated behavior. The non-transitive preferences make it difficult to choose a stable τ.

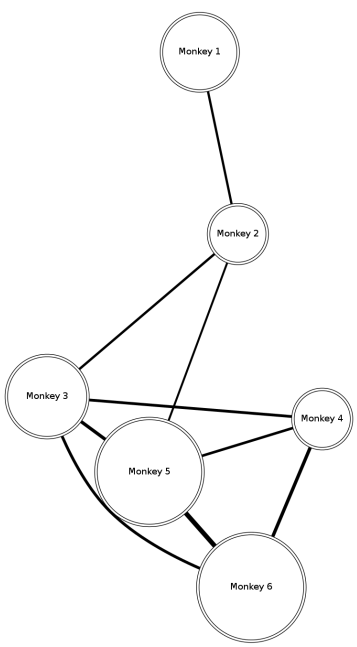


Figure 10: Association preferences for live animals. Diameter of the node is determined by the sum of association preferences for that node. Links are included if they are larger than the mean association preference, and their width is determined by how strong the association preference between the two nodes is.

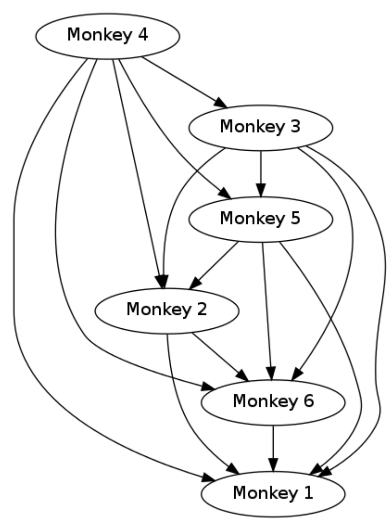


Figure 11: Dominance hierarchy for live animals. Linear chain hierarchies match with our simulated model of dominance behavior, but it is important to note that no part of our inference of dominance relationships *enforces* linear chains. So if the live animals had been an egalitarian troop with little to no aggressive displays, or if the dominance rankings had not been established, we would expect to see a different kind of dominance structure.

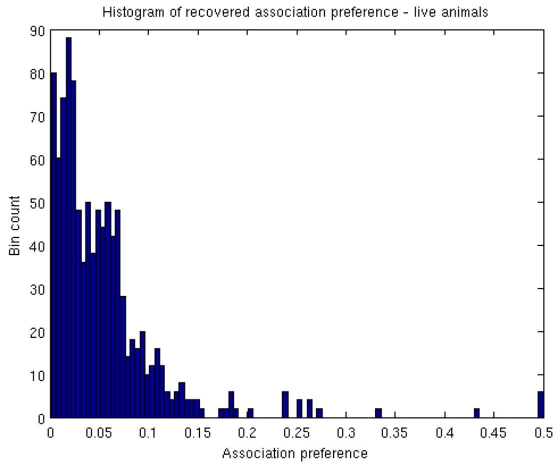


Figure 12: Histogram of association preference values recovered from the live animals. The secondary mode in this distribution is less distinct, but there still is a separation into low and high preference levels.

In future work we would like to expand our approach to include more sophisticated inference techniques. It may be the case that fairly simple probabilistic models will be able to capture some of the structure that we are not able to recover directly from observed behaviors. Association and dominance are clearly not independent relationships, and incorporating some notion of how each affects the other may allow us to improve the accuracy of our inference.

Acknowledgements

This research was supported by the National Institute for Mental Health (MH050268-14S1) as well as by the National Center for Research Resources to the Yerkes National Research Center (P51 RR00165; YNRC Base grant), which is currently supported by the Office of Research Infrastructure Programs/OD P51OD11132. The YNPRC is fully accredited by Americans for the Assessment and Accreditation of Laboratory Care, International.

References

Hemelrijk, C. K. (2000). Towards the integration of social dominance and spatial structure. *Animal Behaviour*, 59(5):1035–1048.

Hrolenok, B. and Balch, T. (2013a). Learning executable models of multiagent behavior from live animal observation. In *ICML 2013 Workshop on Machine Learning For System Identification*.

Hrolenok, B. and Balch, T. (2013b). Learning schooling behavior from observation. In *Advances in Artificial Life, ECAL*, volume 12, page 686691.

Huang, P., Sawhney, R., Walker, D., Wallen, K., Bobick, A., Qin, S., and Balch, T. (2012). Learning a projective mapping to locate animals in video using rfid. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 3830–3836. IEEE.

Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science*, 308(5722):648–652.

Stephens, S. B. and Wallen, K. (2013). Environmental and social influences on neuroendocrine puberty and behavior in macaques and other nonhuman primates. *Hormones and behavior*, 64(2):226–239.

Wallen, K. (1996). Nature needs nurture: the interaction of hormonal and social influences on the development of behavioral sex differences in rhesus monkeys. *Hormones and Behavior*, 30(4):364–378.

Yang, Y.-T., Quitmeyer, A., Hrolenok, B., Shang, H., Nguyen, D. B., Balch, T. R., Medina, T., Sherer, C., and Hybinette, M. (2012). Ant hunt: Towards a validated model of live ant hunting behavior. In *FLAIRS Conference*.