Agent-based simulation for reconstructing social structure by observing collective movements with special reference to single-file movement

Hiroki Koda*[1], Zin Arai†[2], and Ikki Matsuda‡[2,3,4,5]

[1]Primate Research Institute, Kyoto University
[2]Chubu University Academy of Emerging Sciences
[3]Wildlife Research Center of Kyoto University
[4]Japan Monkey Centre
[5]Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah

Abstract

Understanding social organization is fundamental for the analysis of animal societies. In this study, animal single-file movement data —serialized order movements generated by simple bottom-up rules of collective movements— are informative and effective observations for the reconstruction of animal social structures using agent-based models. For simulation, artificial 2-dimensional spatial distributions were prepared with the simple assumption of clustered structures of a group. Animals in the group are either independent or dependent agents. Independent agents distribute spatially independently each one another, while dependent agents distribute depending on the distribution of independent agents. Artificial agent spatial distributions aim to represent clustered structures of agent locations —a coupling of "core" or "keystone" subjects and "subordinate" or "follower" subjects. Collective movements were simulated following two simple rules, 1) initiators of the movement are randomly chosen, and 2) the next moving agent is always the nearest neighbor of the last moving agents, generating "single-file movement" data. Finally, social networks were visualized, and clustered structures reconstructed using a recent major social network analysis (SNA) algorithm, the Louvain algorithm, for rapid unfolding of communities in large networks. Simulations revealed possible reconstruction of clustered social structures using relatively minor observations of single-file movement, suggesting possible application of single-file movement observations for SNA use in field investigations of wild animals.

*koda.hiroki.7a@kyoto-u.ac.jp
†zin@isc.chubu.ac.jp
‡ikki-matsuda@isc.chubu.ac.jp

Significance statements

- An agent-based model was developed to test if animal single-file movement is informative for the reconstruction of clustered social structures.
- Single-file movement data sets were generated with simple bottom-up rules of collective movements.
- Simulations showed possible reconstruction of clustered structures for most combinations of parameter settings.
- Latent structures of clustered animal groups might be identified by collective movement observations, and single-file movement may be practically applied to wildlife investigations.

1 Introduction

Identifying latent structures underlying animal social organization is fundamental in animal ecology [1], as well as in the social sciences [2]. The recent update of computational models for social network analysis (SNA) enables quantitative evaluation of complex processes of social networks based using big data. The basic framework of SNA envisions multiple nodes as animals, viruses, people, or any other agents within the network and at its edges as a dyadic relation mathematically defined between the two nodes [e.g., 3, 4, 5]. The SNA aims to statistically link network structures with bottom-up rules defining dyadic relations, e.g., affiliations, friendships, proximities, co-occurrences, follower/followee, signal sender/receiver, or communications [6]. SNA provides a modern powerful interface for visualizing complex biological networks.

One major biological interest for SNA is statistical approximations of modular, clustered or hierarchical structures embedded in animal social groups [1, 7]. Doubtless, kinship is a basic concept in the theory of the socioecology [for a classical kickoff, e.g., 8]. A mother invests effort in her offspring, mother and offspring generally maintain proximity, and degrees of the social bonding are reflected by relatedness [9, 10]. Based on the cumulative effects of such relatedness, social structures are determined by additional social relations, such as male-female sexual interactions and male-male competition [11, 12]. For example, in primate societies — the most investigated animal society from both theoretical and empirical perspectives— usually display an "alpha" subject, the most dominant social rank, and are composed of the kin-based clusters, such as the female affiliative bonding [13, 14]. Consequently, the "core", "leader" or "influencer" of the animal group are recognized, and multi-layered clustered structures of the group emerge [15, 1, 16, 17]. Some metrics commonly implemented in the modern SNAs provide statistical indices for modularity, centrality or cohesiveness [e.g., 18], and have been applied to actual observational data of animals. These indices support the evaluation of clustered structures [for a recent application in wild chimpanzees, see 19].

For the practical applications of SNA to observational data, dyadic social interactive behaviors or distance-based proximities have been commonly used in ecological analyses of animal societies [3]. Allogrooming is a typical example of dyadic social interaction widely observed in social mammals. Allogrooming is an altruistic behavior and an appropriate behavioral metric to evaluate affiliative interactions [e.g., 20, 21, 22]. Likewise, agonistic interactions, where dominant animals show aggression toward subordinate animals, are also interpretable and countable behaviors to illustrate dominant hierarchy [e.g., 23, 24, 25]. Despite interpretable advantages of social relations, disadvantages of such data are "specializations" of social behaviors —grooming might not always be seen in the insects, or we may not know about the agonistic interactions in some animals— limiting the analysis. In contrast, distance-based data lack contextual information and are not perfectly linked to social relations [26]. These data are generally accepted to indirectly reflect social affiliations [27]. Affiliated individuals may maintain proximity as typically seen in mother-offspring pairs. The recent advanced tracking methods using high-precision GPS [e.g., 28, 29] or Bluetooth beacons [e.g., 30] have strongly influenced location-based proximity metrics due to dramatic expansion of data for interactions among "trackable" animals. Thus, a key analytical step for SNA is selection of appropriate data, which researchers record massively, easily, reliably and precisely.

In this study, another kind of testable data for the SNA is proposed, that is, "animal single-file movement" generated by bottom-up rules in the theory of collective movements. This topic is currently of great interest in animal ecology [for a recent review, e.g., 31]. Typically in fish schools, bird flocks, insect swarms, or ungulate herds, large numbers of individual agents show highly-coordinated collective movements [32, 33]. Theoretically, collective movements do not rely on decision-making of specific individuals, e.g., a "leader", but emerge in a bottom-up manner

such as iterated interactions between animals, either directly (e.g. via visual cues) or indirectly (e.g. via trail formation). Real-time distance-based proximities among all animal agents empirically reveal core mechanisms underlying collective movement emergences. Simultaneous tracking of multiple animals using wearable tags is technologically possible, but perfect tracking of suitable numbers of animals in groups/units for analyses of collective movement are still limited to laboratory-based animals [e.g., 34, 30] or capturable wild animals [35, 36, 37].

Animal single-file movement, defined as serially ordered patterns of animals that are indirectly determined by degrees of inter-agent relations, is a promising target for data collection for linking SNA linking with recent discussions of collective movement by social animals. First, single-file movements are observable in many free-ranging animal species. Classically, social animals moving in single file provided an excellent opportunity to count group members inhabiting dense forests with poor visibility, and further to infer social structures. Extensive historical fieldwork reveals that such events provide information on latent key factors underlying animal social organizations, e.g., dominance hierarchy [38, 39], kin relationships [40], and physiological status, such as breeding cycle [41, 42]. Recent studies focusing on single-file movement and social system structure have examined several animal taxa, e.g., wolves [*Canis lupus*: 38]; zebra [*Equus burchellii*: 41], buffalo [*Syncems cuffer*: 42], chimpanzee [*Pan troglodytes*: 43], mandrill [*Mandrillus sphinx*: 44] and spider monkey [*Ateles geoffroyi*: 45]. Second, animal single-file movement can be recorded without capture and release for attachment of location tags, simply by video-recording (e.g., camera-trappings). This advantage for studying collective movement avoids use of GPS collars that are a burden for free-ranging animals, and most cannot be used for threatened and endangered species [46]. Third, serial order patterns governed by collective movements allow behavioral ecologists to examine individual-based observation essential for documentation and measurement of animal behavior.

This study aimed to determine if movement order in single-file movements are informative and effective observations for the reconstruction of animal social structures using an agent-based computational simulation model [e.g., 47, 48]. An initial artificial 2-dimensional spatial distribution was developed using the simple assumption of clustered group structures. Animal agents in a group are composed of independent and dependent agents. Independent agents distribute spatially independently each other, while dependent agents distribute depending movements of independent agents. The artificial spatial distributions aim to represent clustered structures of agent locations —a coupling of "core" or "keystone" subjects and "subordinate" or "follower" subjects. Collective movements were governed by two simple rules, 1) initiators of the movement are randomly chosen, and 2) the next moving agent is always the nearest neighbor of the last moving agents. These bottom-up rules allowed simulation of "single-file movement" data. Finally, we visualized social networks and reconstructed clustered structures using a recent major SNA algorithm, the Louvain algorithm, for rapid unfolding of communities in large networks [49]. The current idea —that artificial agents are initially generated, collective movements are produced, and collective movements are used to evaluate social networks— is similar to a previous computer experiment [50] that simulated a primate group. The current simulation is more generalized. Parameter assumptions are minimum and used mixed normal distributions. Parameter conditions were sought where single-file movement data work well for reconstruction of the latent clustered structures and further discuss possibilities for practical applications in field investigations of wild animal societies.

2 Methods

The simulations below aim to examine if individual-by-individual serial ordered movements, i.e., "animal single-file movements", provide sufficient information to reconstruct cluster organizations that are typically believed to exist in animal social groups such as primates. Our simulations adopt agent-based models and involve four main steps: 1) animal agents are distributed in 2-dimensional space with latent cluster organizations, 2) those generated agents are serially aligned following a simple collective movement rule (generating the "single-file movements"), 3) then the social network structure is computed using association index defined by serial orders of animal single-file movements, and 4) finally, the cluster organization of the generated social network is evaluated. Figure 3 is a schematic representation of the simulation process. The details of each step are described below.

2.1 Distributing animals (1st step)

First of all, we assume that the animal group of our interest consists of n_I subgroups. Each of these subgroups has the unique *independent* animal agent (*independenter*), who moves independently of any other agents, and n_D *dependent* animal agents (*depender*), whose spatial locations always depend on the independenter of the subgroup. For simplicity, the number n_D of dependers in subgroups is assumed to be the same for all subgroups. Accordingly, the total number of animal agents is $N = n_I(n_D + 1)$. We call the *independenter* of the i-th subgroup the i-th *independenter*.

Locations of independenters are independently generated from the bivariate Gaussian distribution: the location of the i-th independenter in the 2-dimensional space is denoted by $\boldsymbol{x}_i \in \mathbb{R}^2$ and is distributed as

$$\boldsymbol{x}_i \sim \mathcal{N}(\mathbf{0}, \Sigma_I) \tag{1}$$

for $i \in \{1, \cdots n_I\}$. Here $\Sigma_I := \begin{pmatrix} \sigma_I^2 & 0 \\ 0 & \sigma_I^2 \end{pmatrix}$ is the covariance matrix of the distribution (σ_I is the standard deviation of the independenter). The mean of distribution is set to $\mathbf{0}$, the origin of the plane.

Similarly, locations of dependers are generated from the bivariate Gaussian distribution, but with centers depending on the locations of independenters: the location of the j-th depender for i-th subgroup is denoted by $\boldsymbol{x}_k$ where $k = n_I + (i \times j)$ and is distributed as

$$\boldsymbol{x}_k = \boldsymbol{x}_{n_I+(i\times j)} \sim \mathcal{N}(\boldsymbol{x}_i, \Sigma_D), \tag{2}$$

for $i \in \{1, \cdots n_I\}$ and $j \in \{1, \cdots n_D\}$. Here $\Sigma_D := \begin{pmatrix} \sigma_D^2 & 0 \\ 0 & \sigma_D^2 \end{pmatrix}$ is the covariance matrix of the distribution (σ_D is the standard deviation of the depender). The mean parameters set to each independenters' locations, $\boldsymbol{x}_i$. For simplicity, we always assume $\sigma_D = 1$.

Consequently, the spatial distributions of N animal agents in the plane depends on three parameters, n_I, n_D, and σ_I. Figure 1 shows a sample spatial distribution with $n_I = 10$, $n_D = 5$, and $\sigma_I = 100.0$.

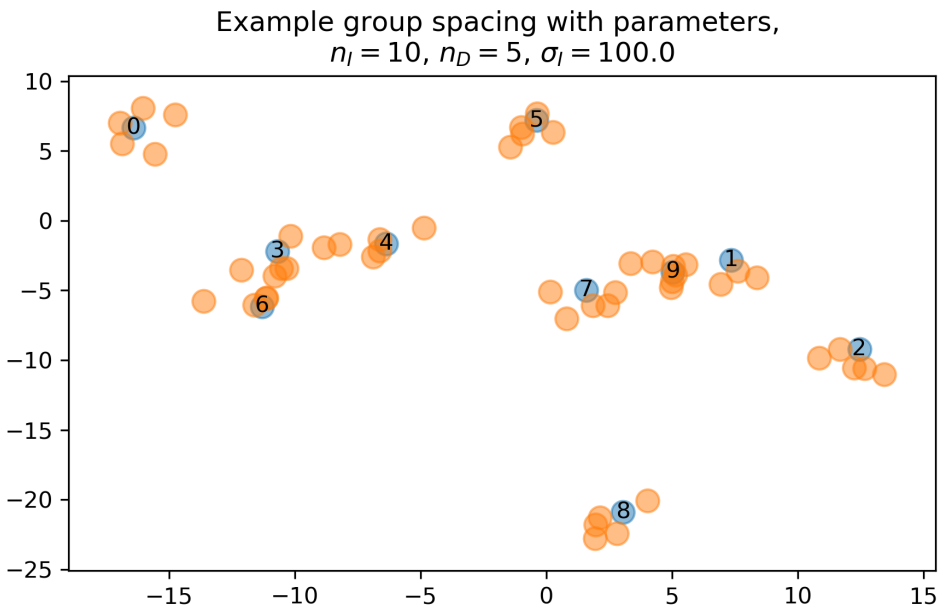

Figure 1: Example of the agent spatial distribution generated with model parameters, $n_I = 10, n_D = 5, \sigma_I = 100.0$.

2.2 Serialization (2nd step)

Subsequently, "animal single-file movement" —the ordered alignment set of animal agents— is generated assuming a simple rule of collective movements defined by the mutual distance between agents. The process is iterative: (1) the agent who first moves (*initiator*) is randomly chosen; (2) in each step of iterations, the nearest agent to the last moved agent is chosen to move next. The order of the single-file movement is defined to be the order of moving in this iterative process.

More specifically, the initiator $\boldsymbol{y}_1$ is randomly chosen from the multi-set of N agents

$$A_1 = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}. \tag{3}$$

Then the next moving agent $\boldsymbol{y}_2$ is chosen by

$$\boldsymbol{y}_2 = \arg\min_{\boldsymbol{x}_k \in A_2} \|\boldsymbol{y}_1 - \boldsymbol{x}_k\| \quad \text{where} \quad A_2 = A_1 \setminus \{\boldsymbol{y}_1\}. \tag{4}$$

Similarly, the i-th moving agent is determined by

$$\boldsymbol{y}_i = \arg\min_{\boldsymbol{x}_k \in A_i} \|\boldsymbol{y}_{i-1} - \boldsymbol{x}_k\| \quad \text{where} \quad A_i = A_{i-1} \setminus \{\boldsymbol{y}_{i-1}\}. \tag{5}$$

Here $\|\cdot\|$ represents the Euclidean distance on the plane. The process of agent selection is repeated $N-1$ times. Consequently, a totally ordered set is generated:

$$C := \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\}, \tag{6}$$

which we call "animal single-file movement". Note that the process of generating C depends on three parameters, n_I, n_D, and σ_I.

We repeat the process of generating the distribution of animals (1st step) and constructing the animal single-file movement (2nd step) n_{exp} times. Here we implicitly mean by n_{exp} the number of observation opportunities for single-file movement in the field. In the following, we therefore assume n_{exp} takes a experimentally feasible value, say, $n_{\text{exp}} = 10$, or 30.

Finally, we denote the set of n_{exp} single-file movements by $\boldsymbol{C} := \{C_1, \ldots, C_{n_{\text{exp}}}\}$.

2.3 Generating social network (3rd step)

In this step, we construct a weighted graph representing the social network structure of animals using the information of $\boldsymbol{C}$.

The graph is defined by an $N \times N$ adjacency matrix $\boldsymbol{G}$ constructed form $\boldsymbol{C}$ as follows. First, given an animal single-file movement C, we define an adjacency matrix G by

$$G = \begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1N} \\ \vdots & \ddots & & & \vdots \\ a_{i1} & & a_{ii} & & a_{iN} \\ \vdots & & & \ddots & \vdots \\ a_{N1} & \cdots & a_{Ni} & \cdots & a_{NN} \end{pmatrix} \tag{7}$$

where $a_{ij} = a_{ji}$ is 1 if the agents $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are adjacent in C, and 0 otherwise (by definition, all diagonal elements are 0). Then the adjacency matrix $\boldsymbol{G}$ for $\boldsymbol{C}$ is defined as

$$\boldsymbol{G} := \sum_{k=1}^{n_{\text{exp}}} G_k \tag{8}$$

where G_k is the adjacency matrix for C_k, the k-th single-file movement in $\boldsymbol{C}$.

For the implementation of this step, we use a recent popular frameworks for social network analysis, `NetworkX` (`https://networkx.github.io/documentation/stable/index.html`) on `python3` and its `from_numpy_matrix()` method (`https://networkx.github.io/documentation/stable/reference/generated/networkx.convert_matrix.from_numpy_matrix.html`). We note that the *parallel_edges=False* option is set; the entries in $\boldsymbol{G}$ are interpreted as the weight of a single edge joining the vertices.

2.4 Clustering (4th step)

Finally, we run a graph clustering algorithm to examine if the subgroup structure of animal agents can be reconstructed from $\boldsymbol{G}$. For this purpose, we apply the Louvain algorithm, one of the greedy optimization algorithms of modularity. The algorithm is implemented in the python module `community` (or `python-louvain` on pypy) as its method `community.best_partition` (see `https://python-louvain.readthedocs.io/en/latest/api.html`). We denote the number of clusters in the result of the Louvain algorithm by n_{eval}. Figure 2 represents an example of the social network graph with clustering obtained by our procedure.

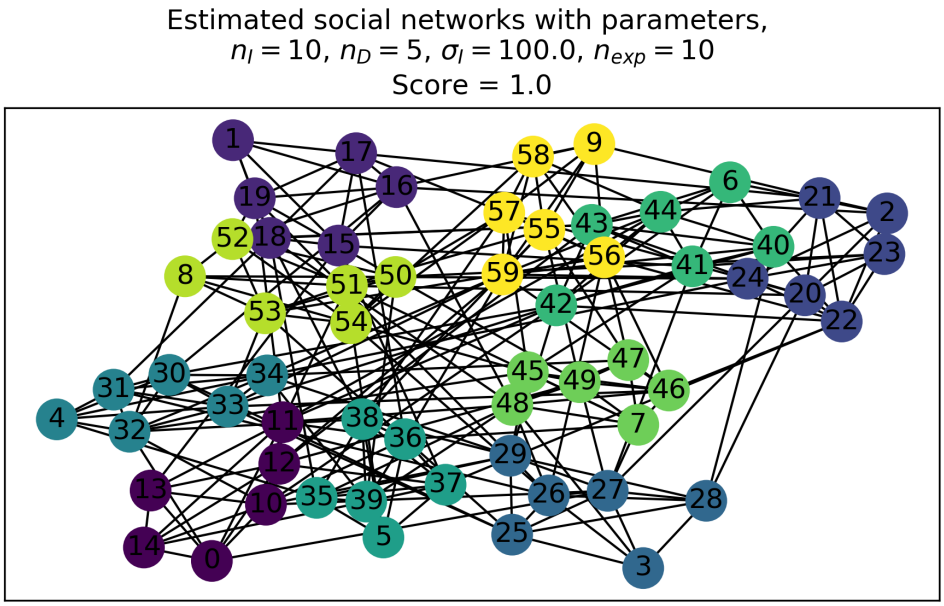

Figure 2: Example of a social network graph clustered by the Louvain algorithm. This graph was generated from the spacing patterns of animal agents shown in Figure 1

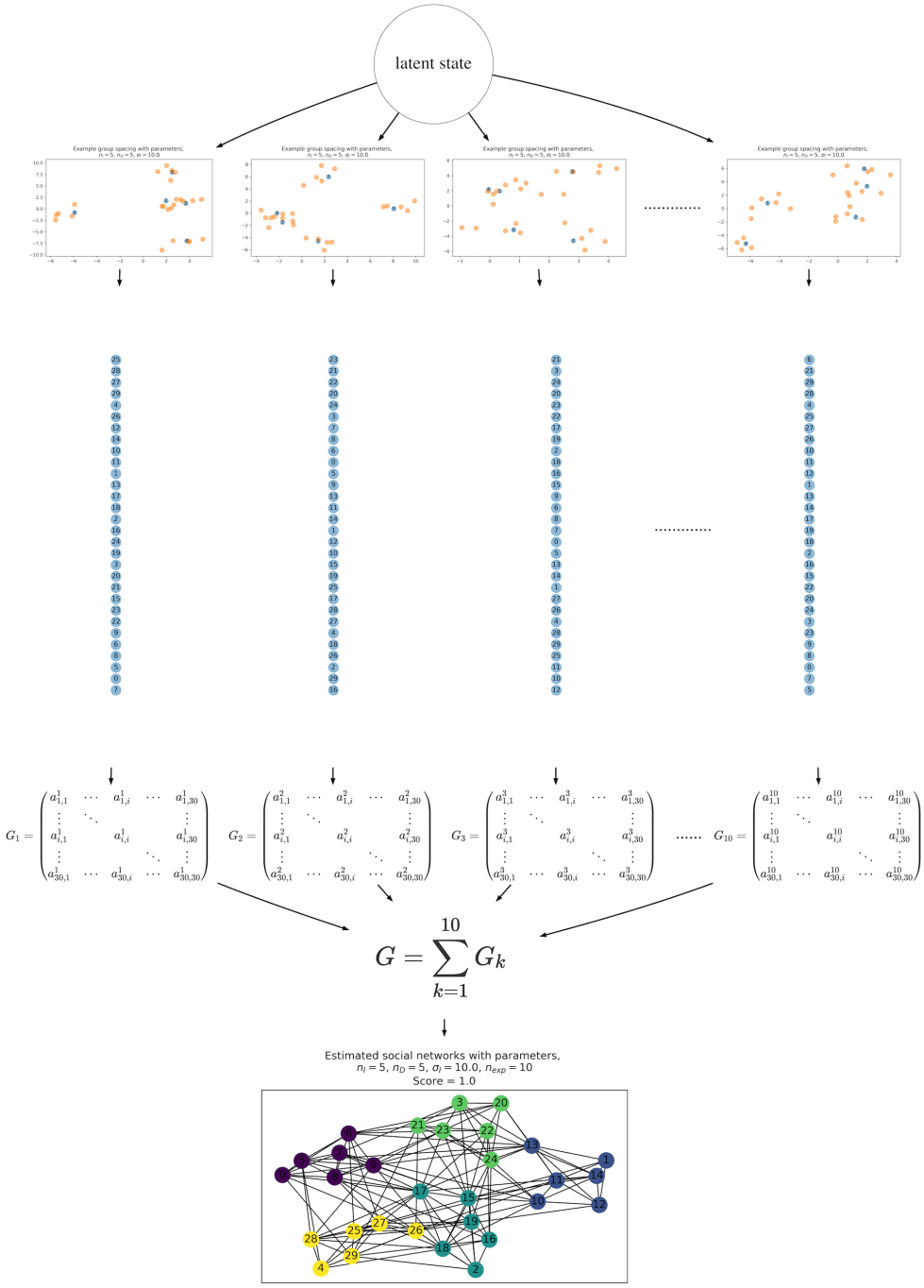

Figure 3: Schematic illustration of the simulation process. In this example, parameters were set as: $n_I = 5, n_D = 5, \sigma_I = 10, \sigma_D = 1, n_{\text{exp}} = 10$. The simulation started from spatial distributions of agents generated from one latent state of social group organization, determined by parameters of the mixed Gaussian processes, $n_I, n_D, \sigma_I, \sigma_D$ (The two top layers). Then, "single-file movement" data sets were generated by the process described in the section on ordered alignment of agents. The vertical chains of the 30 circles represents single-file movement (the number of each circle is the agent id). Adjacency matrices G_k for $k \in \{1, 2, ...10\}$ were generated from orders in the single-file movement data set and were convolved to one adjacency matrix $\boldsymbol{G}$, which was passed to the social network analysis. Finally, the social network graph was produced, with clustering of the local community (bottom graph). This flow is the single simulation process, which is run 1000 times.

2.5 Evaluation of the method

Here we want to estimate how good the clustering we obtained is. For this, we simply count the number of clusters obtained, n_{eval}, and compare it with the "correct answer", the number of subgroups n_I. Namely, we consider the ratio $r = n_{\text{eval}}/n_I$ and call it the score of the cluster estimations. When $r = 1$, our procedure successfully recovers the information of the number of clusters, while $r > 1$ or $r < 1$ implies overestimation or underestimation is taking place, respectively. For each combination of parameters $n_I \in \{1, 2, \cdots 20\}$, $n_D \in \{1, 2, \cdots 20\}$, $n_{\text{exp}} \in \{10, 30\}$, and $\sigma_I \in \{1.0, 2.0, 5.0, 10.0, 100.0\}$, we run the whole procedure 1000 times and compute the average $\bar{r}$ of the score r. The result will be discussed in Section 3.

2.6 A remark on the method

As we will see in Section 3, the proposed method works reasonably fine for most of selected parameter values. In this subsection, we try to explain why this is the case by showing the modularity of the "true" clustering is likely to be higher than that of a clustering obtained by attaching two clusters of the true clustering together. Since the Louvain algorithm optimizes the modularity through a process of inductive amalgamation of clusters, if the true clustering has this property, the algorithm can be expected to end up with the true clustering, which implies the score r will be 1.

We begin by recalling the definition of modularity. Consider a weighted graph $\mathcal{G}$ with n nodes defined by a symmetric $n \times n$ weighted adjacency matrix $G = (a_{ij})$ and a clustering of $\mathcal{G}$ into mutually disjoint sets of nodes $C_1, C_2, \ldots, C_{n_I}$ where n_I is the number of clusters. Then the modularity of this clustering is defined to be

$$Q = \frac{1}{2M} \sum_{i,j=1}^{n} \left(a_{ij} - \frac{k_i k_j}{2M} \right) \delta(C(i), C(j)) \tag{9}$$

where k_i and k_j are the sum of the weights of edges from nodes i and j, respectively; M is the sum of the weight of all edges in the graph; $C(i)$ and $C(j)$ are the clusters to which nodes i and j belong, respectively; δ is the Kronecker delta function. Thus, $\delta(C(i), C(j))$ is 1 if i and j belong the same cluster and 0 otherwise. Note that $2M = \sum_{i,j} a_{ij}$. For $\boldsymbol{G}$ constructed in the 3rd step of our method, since an animal single-file movement has $N - 1$ edges, $M = (N - 1)\, n_{\text{exp}}$ holds.

Let $\{C_1, C_2, \ldots, C_{n_I}\}$ be the true clustering, that is, the clustering exactly reflects the original subgroup structure of animals, and denote its modularity by Q. Then consider a clustering $\{C_1, \ldots, C_p \cup C_q, \ldots, C_{n_I}\}$ obtained from the true clustering by attaching C_p and C_q together and denote its modularity by Q'. A direct calculation shows that

$$Q' - Q = \frac{m}{M} - \frac{K_p K_q}{2M^2} \tag{10}$$

where m is the sum of the weights for all edges from C_p to C_q; K_p and K_q are the sum of the weights for all edges from C_p and C_q, respectively. We want show that this quantity is negative, that is,

$$m < \frac{K_p K_q}{2M}. \tag{11}$$

For simplicity, we assume that in every single-file movement, no mixture of subgruops happens; that is, first $n_D + 1$ agents in the movement belong to the same subgroup, and the next $n_D + 1$ belong to another one, and so on. This is more or less equivalent to assuming that σ_I is very large and therefore each subgroups are completely separated. In this case, (11) can be reasoned as follows. A single-file movement assigns n_D edges to the first and last clusters and

$n_D + 1$ edges to others. Therefore, $K_p, K_q \leq n_D n_{\text{exp}}$ follows. Hence, a sufficient condition for (11) to hold is

$$m < \frac{(n_D n_{\text{exp}})^2}{2M} = \frac{n_D^2 n_{\text{exp}}}{2(n_I(n_D+1)-1)} = \frac{n_D}{2} \cdot \frac{n_D}{(n_D+1) - 1/n_I} \cdot \frac{n_{\text{exp}}}{n_I}. \tag{12}$$

On the other hand, the number of edges between agents that belong to different subgroups is $n_I - 1$ for every single-file movement. Therefore, the sum of the weights for edges in $\mathcal{G}$ connecting different subgroups is $(n_I - 1)\, n_{\text{exp}}$. Thus the expected value of m is

$$\frac{(n_I - 1)\, n_{\text{exp}}}{n_I(n_I - 1)} = \frac{n_{\text{exp}}}{n_I} \tag{13}$$

provided that all the ordering of subgroups in single-file movements are equally probable. The right hand side of (12) can now be understood as the expected value of m multiplied by

$$\frac{n_D}{2} \cdot \frac{n_D}{(n_D+1) - 1/n_I}, \tag{14}$$

which is approximately $n_D/2$ in our parameter settings. This implies that, when n_{exp} is large enough, (12) is very likely to hold for all p, q.

3 Experimental Results

Figure 4 is the contour plots of the calculated average score $\bar{r}$ for some selected sets of parameter values. This suggests that the average score increases as parameters σ_I and n_{exp} increases, just as expected. It is also observed that the score decreases as n_I, the numbers of the independenter, increases. In contrast, it increases as n_D, the numbers of the depender, increases. In summary, scores will be better with large $\sigma_I, n_{\text{exp}}, n_D$ and small n_I (expect for the case of $n_I = 1$). In the cases of $n_{\text{exp}} = 30$ and $\sigma_I \geq 5$, scores were often perfect, i.e., $\bar{r} \approx 1$. Thus, cluster size estimations were often successful, if animal single-file movements were observed 30 times.

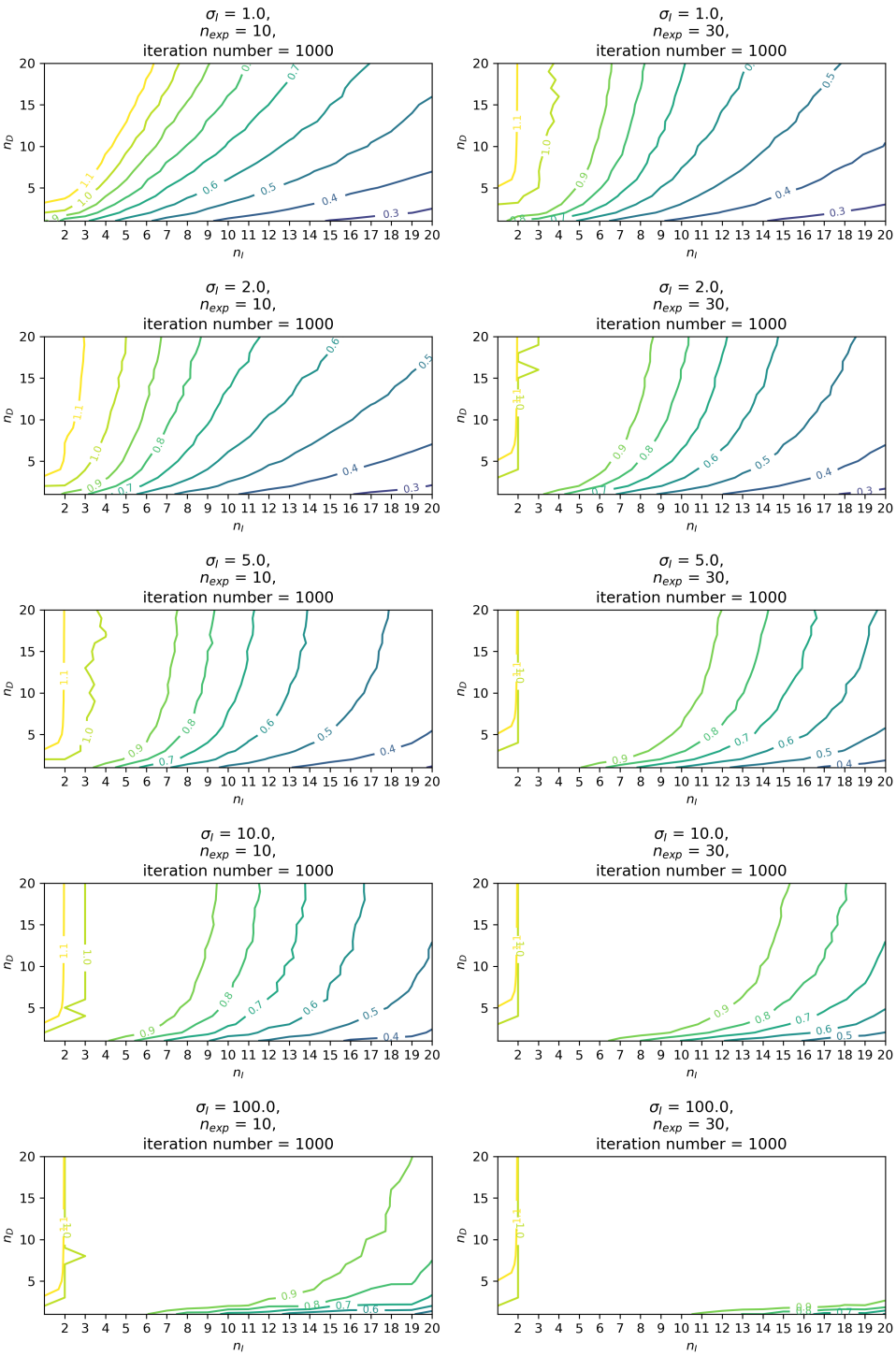

Figure 4: Contour plots for average performance scores for cluster estimation, $\bar{r}$ in 1000-steps simulations. Areas near the score of 1.0 represents areas of parameter combinations where perfect estimation scores returned.

4 Discussion

Simulations with agent-based models suggest that recent implementation of a community clustering algorithm in the SNA could well estimate latent community structures generated by mixed Gaussian distributions under several parameter conditions. First, SD ratio parameters, $\frac{\sigma_I}{\sigma_D}$, likely influence performance for identification of clustering structures. The greater these ratios are, the more precise the estimations of cluster size obtained. This result is reasonable since the contrast of the modular structure becomes clear (see Figure 1 as an example of the $\sigma_I = 100$). Likewise, the more iterations, the more accurate scores of clusters size became. This result is also not surprising. An important result is the effect of numbers of *dependers* ffor the cluster estimation. Results consistently indicate that greater numbers of *dependers* improved estimation performance, at least within the range of parameter settings used in study simulations. Too many *dependers* would likely lead to overestimations of the cluster sizes, but such confusion occurred only in the case of $\sigma_I = 1, n_{\text{exp}} = 10$, and did not occur using most other parameter combinations.

An additional lesson from the simulations is an implication that appropriate numbers of the *independenters* exixt for better performance in cluster estimation. At least for parameter settings of $1 \leq n_I \leq 30$, a single-digit number, except for 1 and 2, likely shows better performance. In summary, results suggested that SNA applications of the Louvain algorithm to single-file movement data work well even when using only 30 iterations in simulations ($n_{\text{exp}} = 30$), when 1) the *independenters* range between around 3 and 10, 2) *dependenters* count is as large as reasonable, and 3) SD ratios of *independenter* per *dependenter* are high.

Looking back to observations of actual animal societies in reality, animals would fit with conditions appropriate for fine reconstruction using the present method? A promising candidate, for example, are primate societies. Primates are middle to large mammals mainly inhabiting tropical forests and sometimes adapting to dry savanna, temperate forests or cool-temperate climates [51]. Primates show a large diversity of social systems [52], from simple monogamous to complex multi-level societies that presumably emerged based on the ecological constrains, such as competition for food and threats of predation, as discussed in successful ecological models [e.g., 53, 54, 14].In general, principles of primate social formation rely on kin-based relationships and due to ecological constrains, group size is typically limited to about 100 individuals, but mostly groups range from 2.5-50 individuals [55]. Group size and structure, which are often consist of a few matrilineal clustered families and a few independent males, seemingly match with the successful parameter conditions indicated in the simulations.

Another advantage of primates is methodological characteristics of primate behavioral ecology. Primates are typically long-lived animals with slow life histories, and long-term studies that have been ongoing continuously for approximately a decade or more, are relatively common [56, 57]. Traditionally, primatologists have made individual long-term observations. Consequently, all group members can be perfectly identified in both field and captive studies. Observations at road or river crossings —a suitable event for single-file movement— have been reported in field observations for several primate species, e.g., chimpanzee [58], mandrill [44] and proboscis monkeys [59]. Such events (single-file movement) would potentially happen frequently in the field, though these observations may be used only for counting group members. Relevant findings of simulations are that the SNA works well for only 30 iterations of the single-file observation, and such observations may well exist for well-studied primates. Given advantages in primate groups, applications of the present SNA to their single-file movement data can be empirically tested soon. The quest to find the social structures of the primates is a fundamental issue, studied by many pioneer primatologists. Thus, the proposed approach provides a novel and empirical tool to analyze primate society, especially for their socio-spatial organization.

The results of the study might also be applied to other unusually social animals. Due to the recent human imposed threats such as roads, animals are often forced into single-file formation to cross risky or dangerous places. Such an aspect of socio-spatial organization under highly risky situations has also been analyzed in meerkat [60] and elephant [61]. Also, single-file movement is naturally observed in zebra heading to water holes [41] and in wolves through snow [62]. Anywhere such single-file movements are often observable recently developed camera-trapping techniques [63] would enable collection of a valuable data set. Further, where animal movement is controllable in semi-free range conditions such as narrow gateways to feeders and water, another opportunity exists. Such single-file movement has been recorded for sheep [*Ovis aries*: 39], and would be ideal to test simulation results in various situations.

Finally, we also acknowledge possible integration with other social network data. The output from the single-file movement-SNA is just one dimension of inference for social structures. For a deep understanding of social organizations, investigation of animal society with multidimensional views and data is needed. For example, genetic test of relatedness is a direct evaluation of network patterns. Behavioral observations, e.g., traditional quantitative grooming counts of the groomer-groomee or vocal exchanges of callers and responders, are also direct metrics that reflect social networks. Gut bacterial flora of groups of individuals was recently suggested as a possible candidate to elucidate the network structures of animal society. Thus, access to many kinds of elemental data is available for predictions concerning social networks. Together with comparisons among network structures reconstructed by different metrics, an ultimate understanding of fundamental mechanisms underlying the animal society will eventually emergence.

Acknowledgements

We appreciate the profound comments and suggestions from Ichiro Tsuda, Aru Toyoda, and Takashi Morita. This study was performed under the Cooperative Research Program at KUPRI (2018-C-27, 2019-B-27)

Funding

This study was mainly funded by the Japan Science and Technology Agency, Core Research for Evolutional Science and Technology 17941861 (#JPMJCR17A4) and partly by MEXT Grant-in-aid for Scientific Research on Innovative Areas #4903 (Evolinguistics), 17H06380.

Code accessibility

All code is available from `https://github.com/hkoda/mix_gauss_caravan_sna`.

Author contributions

Project organization: HK IM; Computational simulations: HK; Analytic discussion: ZA; Manuscript writing: HK ZA IM.

References

[1] Amiyaal Ilany and Erol Akçay. Social inheritance can explain the structure of animal social networks. *Nature Communications*, 7(1):12084, nov 2016.

[2] David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Albert László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. Social science: Computational social science. *Science*, 323(5915):721–723, 2009.

[3] Darren P. Croft, Richard James, and Jens Krause. *Exploring animal social networks*. Princeton University Press, 2008.

[4] Ivan Puga-Gonzalez, Sebastian Sosa, and Cédric Sueur. Editorial: Social networks analyses in primates, a multilevel perspective. *Primates*, 60(3):163–165, may 2019.

[5] Cédric Sueur, Jean-Louis Deneubourg, Odile Petit, and Iain D. Couzin. Group size, grooming and fission in primates: A modeling approach based on group structure. *Journal of Theoretical Biology*, 273(1):156–166, mar 2011.

[6] Jens Krause, Richard James, Daniel W Franks, and Darren P Croft. *Animal Social Networks*. Oxford University Press, USA, 2015.

[7] Cristian Pasquaretta, Marine Levé, Nicolas Claidière, Erica van de Waal, Andrew Whiten, Andrew J. J. MacIntosh, Marie Pelé, Mackenzie L. Bergstrom, Christèle Borgeaud, Sarah F. Brosnan, Margaret C. Crofoot, Linda M. Fedigan, Claudia Fichtel, Lydia M. Hopper, Mary Catherine Mareno, Odile Petit, Anna Viktoria Schnoell, Eugenia Polizzi di Sorrentino, Bernard Thierry, Barbara Tiddi, and Cédric Sueur. Social networks in primates: smart and tolerant species have more efficient networks. *Scientific Reports*, 4(1):7600, may 2015.

[8] Robert L. Trivers. The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, 46(1):35–57, mar 1971.

[9] Joan B. Silk, Susan C. Alberts, and Jeanne Altmann. Social Bonds of Female Baboons Enhance Infant Survival. *Science*, 302(5648):1231–1234, nov 2003.

[10] Joan B. Silk, Jacinta C. Beehner, Thore J. Bergman, Catherine Crockford, Anne L. Engh, Liza R. Moscovice, Roman M. Wittig, Robert M. Seyfarth, and Dorothy L. Cheney. Strong and consistent social bonds enhance the longevity of female baboons. *Current Biology*, 20(15):1359–1361, aug 2010.

[11] Tim Clutton-Brock. *Mammal Societies*. John Wiley & Sons, 2016.

[12] Tim Clutton-Brock and Charles Janson. Primate socioecology at the crossroads: Past, present, and future. *Evolutionary Anthropology*, 21(4):136–150, jul 2012.

[13] Susanne Shultz and Robin Dunbar. Bondedness and sociality. *Behaviour*, 147(7):775–803, 2010.

[14] Richard W Wrangham. An Ecological Model of Female-Bonded Primate Groups. *Behaviour*, 75(3-4):262–300, 1980.

[15] A. Baigger, N. Perony, M. Reuter, V. Leinert, M. Melber, S. Grünberger, D. Fleischmann, and G. Kerth. Bechstein's bats maintain individual social links despite a complete reorganisation of their colony structure. *Naturwissenschaften*, 100(9):895–898, sep 2013.

[16] Xiao-Guang Qi, Paul A. Garber, Weihong Ji, Zhi-Pang Huang, Kang Huang, Peng Zhang, Song-Tao Guo, Xiao-Wei Wang, Gang He, Pei Zhang, and Bao-Guo Li. Satellite telemetry and social modeling offer new insights into the origin of primate multilevel societies. *Nature Communications*, 5(1):5296, dec 2014.

[17] Sara B. Tavares, Filipa I.P. Samarra, and Patrick J.O. Miller. A multilevel society of herring-eating killer whales indicates adaptation to prey characteristics. *Behavioral Ecology*, 28(2):500–514, apr 2017.

[18] Aric A Hagberg, Daniel A Schult, and Pieter J Swart. Exploring Network Structure, Dynamics, and Function using NetworkX. In Gaël Varoquaux, Travis Vaught, and Jarrod Millman, editors, *Proceedings of the 7th Python in Science Conference*, pages 11–15, Pasadena, CA USA, 2008.

[19] Daniel Schofield, Arsha Nagrani, Andrew Zisserman, Misato Hayashi, Tetsuro Matsuzawa, Dora Biro, and Susana Carvalho. Chimpanzee face recognition from videos in the wild using deep learning. *Science Advances*, 5(9):eaaw0736, sep 2019.

[20] Patricia Kanngiesser, Cédric Sueur, Katrin Riedl, Johannes Grossmann, and Josep Call. Grooming network cohesion and the role of individuals in a captive chimpanzee group. *American Journal of Primatology*, 73(8):758–767, aug 2011.

[21] Marine Levé, Cédric Sueur, Odile Petit, Tetsuro Matsuzawa, and Satoshi Hirata. Social grooming network in captive chimpanzees: does the wild or captive origin of group members affect sociality? *Primates*, 57(1):73–82, jan 2016.

[22] Ikki Matsuda, Keiichi Fukaya, Cristian Pasquaretta, and Cédric Sueur. Factors influencing grooming social networks: insights from comparisons of colobines with different dispersal patterns. In *Dispersing Primate Females*, pages 231–254. Springer, 2015.

[23] Jake A. Funkhouser, Jessica A. Mayhew, and John B. Mulcahy. Social network and dominance hierarchy analyses at Chimpanzee Sanctuary Northwest. *PLoS ONE*, 13(2):e0191898, feb 2018.

[24] Tatsuro Kawazoe and Sebastian Sosa. Social networks predict immigration success in wild Japanese macaques. *Primates*, 60(3):213–222, may 2019.

[25] Julia Lehmann, Bonaventura Majolo, and Richard McFarland. The effects of social network position on the survival of wild Barbary macaques, Macaca sylvanus. *Behavioral Ecology*, 27(1):20–28, 2016.

[26] Madelaine Castles, Robert Heinsohn, Harry H. Marshall, Alexander E.G. Lee, Guy Cowlishaw, and Alecia J. Carter. Social networks created with different techniques are not comparable. *Animal Behaviour*, 96:59–67, oct 2014.

[27] Damien R. Farine. Proximity as a proxy for interactions: Issues of scale in social network analysis. *Animal Behaviour*, 104:e1–e5, jun 2015.

[28] Gaëlle Fehlmann and Andrew J. King. Bio-logging. *Current Biology*, 26(18):R830–R831, sep 2016.

[29] Christian Rutz and Graeme C. Hays. New frontiers in biologging science. *Biology Letters*, 5(3):289–292, jun 2009.

[30] Takashi Morita, Aru Toyoda, Seitaro Aisu, Akihisa Kaneko, Naoko Suda-Hashimoto, Ikki Matsuda, and Hiroki Koda. Animals exhibit consistent individual differences in their movement: A case study on location trajectories of Japanese macaques. *Ecological Informatics*, 56(January):101057, mar 2020.

[31] Peter A. H. Westley, Andrew M. Berdahl, Colin J. Torney, and Dora Biro. Collective movement in ecology: from emerging technologies to conservation and management. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746):20170004, may 2018.

[32] Iain D. Couzin and Jens Krause. Self-Organization and Collective Behavior in Vertebrates. In *Advances in the Study of Behavior*, volume 32, pages 1–75. 2003.

[33] Cédric Sueur and Jean-Louis Deneubourg. Self-Organization in Primates: Understanding the Rules Underlying Collective Movements. *International Journal of Primatology*, 32(6):1413–1432, dec 2011.

[34] Franziska Boenisch, Benjamin Rosemann, Benjamin Wild, David Dormagen, Fernando Wario, and Tim Landgraf. Tracking All Members of a Honey Bee Colony Over Their Lifetime Using Learned Models of Correspondence. *Frontiers in Robotics and AI*, 5(APR):1–10, apr 2018.

[35] Frédéric Lesmerises, Chris J. Johnson, and Martin Hugues St-Laurent. Landscape knowledge is an important driver of the fission dynamics of an alpine ungulate. *Animal Behaviour*, 140:39–47, jun 2018.

[36] Marie Sigaud, Jerod A. Merkle, Seth G. Cherry, John M. Fryxell, Andrew Berdahl, and Daniel Fortin. Collective decision-making promotes fitness loss in a fusion-fission society. *Ecology Letters*, 20(1):33–40, jan 2017.

[37] A. Strandburg-Peshkin, D. R. Farine, I. D. Couzin, and M. C. Crofoot. Shared decision-making drives collective movement in wild baboons. *Science*, 348(6241):1358–1361, jun 2015.

[38] Rolf O Peterson, Amy K Jacobs, Thomas D Drummer, L David Mech, and Douglas W Smith. Leadership behavior in relation to dominance and reproductive status in gray wolves, Canis lupus. *Canadian Journal of Zoology*, 80(8):1405–1412, aug 2002.

[39] V.R. Squires and G.T. Daws. Leadership and dominance relationships in Merino and Border Leicester sheep. *Applied Animal Ethology*, 1(3):263–274, jul 1975.

[40] John C. Mitani, D.Andrew Merriwether, and Chunbin Zhang. Male affiliation, cooperation and kinship in wild chimpanzees. *Animal Behaviour*, 59(4):885–893, apr 2000.

[41] Ilya R. Fischhoff, Siva R. Sundaresan, Justine Cordingley, Heather M. Larkin, Marie-Jeanne Sellier, and Daniel I. Rubenstein. Social relationships and reproductive state influence leadership roles in movements of plains zebra, Equus burchellii. *Animal Behaviour*, 73(5):825–831, may 2007.

[42] H. H.T. Prins. Buffalo Herd Structure and its Repercussions for Condition of Individual African Buffalo Cows. *Ethology*, 81(1):47–71, apr 1989.

[43] John C. Mitani and David P. Watts. Correlates of territorial boundary patrol behaviour in wild chimpanzees. *Animal Behaviour*, 70(5):1079–1086, nov 2005.

[44] Shun Hongo. New evidence from observations of progressions of mandrills (Mandrillus sphinx): a multilevel or non-nested society? *Primates*, 55(4):473–481, oct 2014.

[45] Filippo Aureli, Colleen M. Schaffner, Jan Verpooten, Kathryn Slater, and Gabriel Ramos-Fernandez. Raiding parties of male spider monkeys: Insights into human warfare? *American Journal of Physical Anthropology*, 131(4):486–497, dec 2006.

[46] Kerry M. Dore, Malene F. Hansen, Amy R. Klegarth, Claudia Fichtel, Flávia Koch, Andrea Springer, Peter Kappeler, Joyce A. Parga, Tatyana Humle, Christelle Colin, Estelle Raballand, Zhi Pang Huang, Xiao Guang Qi, Anthony Di Fiore, Andrés Link, Pablo R. Stevenson, Danica J. Stark, Noeleen Tan, Christa A. Gallagher, C. Jane Anderson, Christina J. Campbell, Marina Kenyon, Paula Pebsworth, David Sprague, Lisa Jones-Engel, and Agustín Fuentes. Review of GPS collar deployments and performance on nonhuman primates. *Primates*, jan 2020.

[47] E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(Supplement 3):7280–7287, may 2002.

[48] Volker Grimm. Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science*, 310(5750):987–991, nov 2005.

[49] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, oct 2008.

[50] Cédric Sueur, Jean-Louis Deneubourg, and Odile Petit. From Social Network (Centralized vs. Decentralized) to Collective Decision-Making (Unshared vs. Shared Consensus). *PLoS ONE*, 7(2):e32566, feb 2012.

[51] Russell A Mittermeier, Don E Wilson, and Anthony B Rylands. *Handbook of the mammals of the world: primates*. Lynx Edicions, 2013.

[52] John C Mitani, Josep Call, Peter M Kappeler, Ryne A Palombit, and Joan B Silk. *The evolution of primate societies*. University of Chicago Press, 2012.

[53] Elisabeth H.M. Sterck, David P. Watts, and Carel P. Van Schaik. The evolution of female social relationships in nonhuman primates. *Behavioral Ecology and Sociobiology*, 41(5):291–309, nov 1997.

[54] C. P. Van Schaik. Why are Diurnal Primates Living in Groups? *Behaviour*, 87(1-2):120–144, 1983.

[55] R. I.M. Dunbar, Padraig Mac Carron, and Susanne Shultz. Primate social group sizes exhibit a regular scaling pattern with natural attractors. *Biology Letters*, 14(1):20170490, jan 2018.

[56] Colin A. Chapman, Amélie Corriveau, Valérie A.M. Schoof, Dennis Twinomugisha, and Kim Valenta. Long-term simian research sites: Significance for theory and conservation. *Journal of Mammalogy*, 98(3):652–660, may 2017.

[57] Peter M Kappeler and David P Watts. *Long-term field studies of primates*. Springer Science & Business Media, 2012.

[58] Kimberley Jane Hockings. Behavioral Flexibility and Division of Roles in Chimpanzee Road-Crossing. In Tetsuro Matsuzawa, Tatyana Humle, and Yukimaru Sugiyama, editors, *The Chimpanzees of Bossou and Nimba*, Primatology Monographs, pages 221–229. Springer Japan, Tokyo, 2011.

[59] Ikki Matsuda, Augustine Tuuga, Yoshihiro Akiyama, and Seigo Higashi. Selection of river crossing location and sleeping site by proboscis monkeys (Nasalis larvatus) in Sabah, Malaysia. *American Journal of Primatology*, 70(11):1097–1101, nov 2008.

[60] Nicolas Perony and Simon W. Townsend. Why Did the Meerkat Cross the Road? Flexible Adaptation of Phylogenetically-Old Behavioural Strategies to Modern-Day Threats. *PLoS ONE*, 8(2):e52834, feb 2013.

[61] Kaori Mizuno, Nachiketha Sharma, Gen'ichi Idani, and Raman Sukumar. Collective behaviour of wild Asian elephants in risky situations: how do social groups cross roads? *Behaviour*, 154(12):1215–1237, 2017.

[62] Rolf Olin Peterson. *Wolf ecology and prey relationships on Isle Royale*. Number 11. Department of the Interior, National Park Service, 1977.

[63] Allan F O'Connell, James D Nichols, and K Ullas Karanth. *Camera traps in animal ecology: methods and analyses*. Springer Science & Business Media, 2010.