

# Ranking Network of a Captive Rhesus Macaque Society: A Sophisticated Corporative Kingdom

Hsieh Fushing<sup>1</sup>, Michael P. McAssey<sup>2</sup>, Brianne Beisner<sup>3</sup>, Brenda McCowan<sup>4\*</sup>

**1** Department of Statistics, University of California Davis, Davis, California, United States of America, **2** Department of Statistics, University of California Davis, Davis, California, United States of America, **3** California National Primate Research Center, University of California Davis, Davis, California, United States of America, **4** California National Primate Research Center, Department of Population Health and Reproduction, School of Veterinary Medicine, University of California Davis, Davis, California, United States of America

## Abstract

We develop a three-step computing approach to explore a hierarchical ranking network for a society of captive rhesus macaques. The computed network is sufficiently informative to address the question: Is the ranking network for a rhesus macaque society more like a kingdom or a corporation? Our computations are based on a three-step approach. These steps are devised to deal with the tremendous challenges stemming from the transitivity of dominance as a necessary constraint on the ranking relations among all individual macaques, and the very high sampling heterogeneity in the behavioral conflict data. The first step simultaneously infers the ranking potentials among all network members, which requires accommodation of heterogeneous measurement error inherent in behavioral data. Our second step estimates the social rank for all individuals by minimizing the network-wide errors in the ranking potentials. The third step provides a way to compute confidence bounds for selected empirical features in the social ranking. We apply this approach to two sets of conflict data pertaining to two captive societies of adult rhesus macaques. The resultant ranking network for each society is found to be a sophisticated mixture of both a kingdom and a corporation. Also, for validation purposes, we reanalyze conflict data from twenty longhorn sheep and demonstrate that our three-step approach is capable of correctly computing a ranking network by eliminating all ranking error.

**Citation:** Fushing H, McAssey MP, Beisner B, McCowan B (2011) Ranking Network of a Captive Rhesus Macaque Society: A Sophisticated Corporative Kingdom. PLoS ONE 6(3): e17817. doi:10.1371/journal.pone.0017817

**Editor:** Matjaz Perc, University of Maribor, Slovenia

**Received:** December 2, 2010; **Accepted:** February 10, 2011; **Published:** March 15, 2011

**Copyright:** © 2011 Fushing et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The first two authors are supported in part by the NSF under Grant No. HSD 0826844 and DMS-1007219, and by the NIH under Grant No. 1R01AG025218-01A2. The last two authors are supported in part by the NIH under Grant Nos. R24 RR024396 (BM) and PR51 RR000169 (CNPRC Base Grant). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: bjmcowan@ucdavis.edu

## Introduction

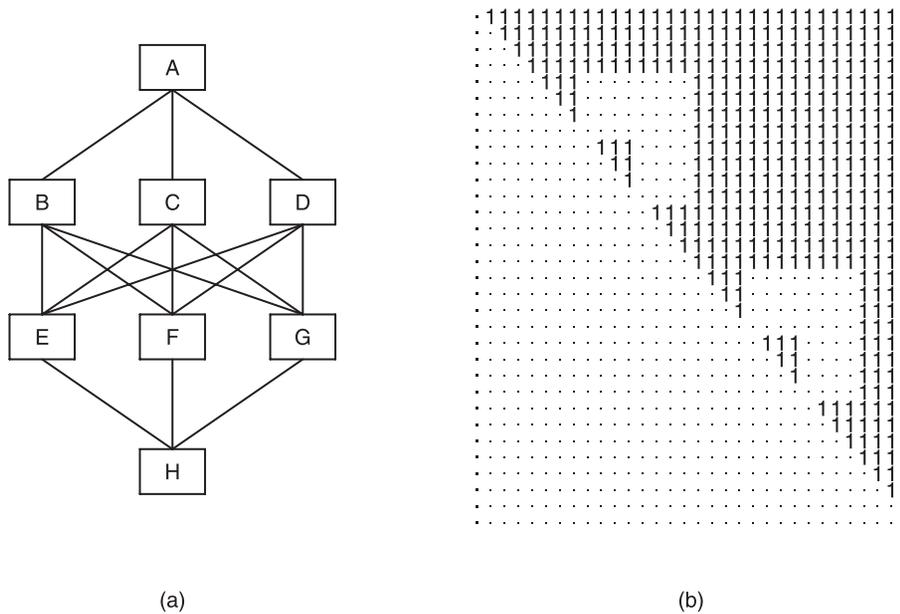
In many animal societies, individual members are understood to possess an intrinsic characteristic called *dominance* which both regulates and is molded by their interactions with other members [1]. Dominance hierarchies are thus a common feature of many such societies, ranging from wasps [2] to crayfish [3] to pigeons [4] to longhorn sheep [5,6]. Furthermore, animal societies have typically been placed into one of two categories: those with defined dominance relationships and those with undefined, or egalitarian, relationships, e.g., [7]. In social groups with detectable dominance relationships, group members are usually arranged into a sequential (often called *linear*) hierarchy based upon observations of direct conflicts between group members. In a sequential hierarchy, the top-ranked member of the group dominates all other members, the second-ranked member dominates all other members except the top-ranked member, and so on. An abundance of mathematical methods have been presented to determine the sequential hierarchy in animal societies based on observed interactions among their members, e.g., [2,6,8,9].

However, in some social groups the dominance hierarchy may not follow a completely sequential pattern, particularly in larger groups and those with highly complex social and cognitive abilities. In addition, the social hierarchy may exist along a

continuum between a sequential dominance hierarchy and an egalitarian society where some group members have clear dominance relationships while others do not [2,9]. Therefore, alternative methods for analyzing the ranking structure of social groups must be sought.

A common example of a non-sequential network is the corporative one. A corporative ranking network allows several groups of individuals to have no dominance relations among them. Typically such a network consists of several tiers: one “governing” group sitting on the top tier, and multiple lower tiers, each of which may contain several parallel groups. Sequential dominance relationships are only found among individuals within the same group and among groups with a group-dominance relation. This ranking network is common in human society, but has rarely been studied using social network analysis [10]. An illustrative network is given in panel (a) of Figure 1. A special case of a corporative ranking network is the kingdom ranking network, in which each group is identified by a common biological linkage, such as matriline or common genetic descent. In particular, the “governing” group is the so-called “royal” group, and royal group members are absolutely not dominated by any non-royal group member in the society.

A distinctive feature of a non-sequential ranking network in general is that transitivity, by which a dominance relationship



**Figure 1. An illustrative matrix representation of an ideal corporative ranking structure.** Panel (a) depicts a hierarchy of four tiers with eight groups. Groups on the same tier have the same rank, Panel (b) shows an upper-triangular matrix consisting mostly of 1 s, with several rectangular patches consisting of 0 s (represented by dots), representing the presence or absence of pairwise dominance relationships among the members of the eight groups.  
doi:10.1371/journal.pone.0017817.g001

between two individuals may be inferred based upon their observed interactions with a common third party, plays a more influential role than that in a sequential ranking network. For instance, if *A* dominates *B*, and *B* dominates *C*, then it is inferred that *A* dominates *C*. With *B* as an intermediate individual, the dominance relationship between *A* and *C* can hold even without observation of direct interactions between *A* and *C*. Thus, in a corporative ranking network, many network parameters, such as dominance potentials, are highly interdependent and interact in a non-sequential fashion. The task of constructing such a complex network, therefore, becomes computationally challenging.

The pertinent challenge can be perceived from two successive levels: The first level involves simultaneous inference of all possible dominance potentials among network members from observed behavioral interactions. These behavioral data involve pairs of network members with and without observed direct interaction, and thus the task of estimating the dominance potentials of all possible pairs is subject to highly heterogeneous measurement error. The second level involves identifying each individual's network coordinate, by arranging all pairwise dominance potentials into an upper-triangular matrix form similar to the illustrating example presented in Figure 1(b).

In the matrix shown in Figure 1(b), a 1 in cell  $(i, j)$  indicates that the *i*th individual dominates the *j*th individual in the society, while a dot indicates otherwise. The rows (and columns) are arranged in such a way that the first four correspond to the members of group *A* at the top of the structure shown in panel (a) of Figure 1, followed by four rows corresponding to the members of group *B*, then group *C*, and so on, with the final four rows corresponding to the members of group *H*. The rectangular patches of dots within the upper triangle occur because the pairwise dominance between members of different groups occupying the same tier of the hierarchy, such as groups *B*, *C* and *D*, cannot be distinguished. However, within any group the members may follow a dominance order, which is indicated by the consecutive small triangular

clusters of 1 s along the diagonal of the matrix. The absence of 1 s in the lower triangle shows that a permutation of the members of this society has been identified such that, if the members are assigned ranks in this order, then no individual of lower rank dominates any individual of higher rank.

Let  $p_{ij}$  denote the dominance potential of group member *i* over member *j*, with  $p_{ji} = 1 - p_{ij}$ . Then *i* dominates *j* if and only if  $p_{ij} > 0.5$ . Hence, given a matrix  $P = [p_{ij}]$  with the rows and columns to be ordered by rank, the goal is to find a permutation of all individuals in the society such that the sum of the dominance potentials exceeding 0.5 in the lower triangle of the matrix are as small as possible. The individual ranks of each network member must be simultaneously estimated based on both direct observation of dominance behavior and inferred dominance due to the assumed transitivity. Thus the challenge becomes a high-dimensional mathematical optimization problem subject to heterogeneous uncertainty. Moreover, since in a non-sequential ranking network many pairs will have dominance potentials near 0.5, more than one set of ranking coordinates can produce a legitimate optimal solution.

The fact that non-sequential ranking network structure has not yet been well-studied in the literature is partly due to the unavailability of reliable data. However, the availability of reliable data has changed dramatically in the past two decades as the awareness of social network theory, aided by information advances, has motivated researchers to collect network-like data, especially in animal behavioral studies. Reliable behavioral data sets of non-human primate societies are available from many institutes across many research fields, such as the California National Primate Research Center (CNPRC).

In this paper we resolve the aforementioned challenges computationally. Three steps are devised. The first step simultaneously estimates all dominance potentials based on both direct pairwise interactions and transitive dominance inferred from interactions with common third parties. Our second step estimates

the ranking coordinates of all individuals by minimizing the cumulative error in the inferred dominance potentials across the entire network. The third step derives confidence bounds for desired empirical features of the computed ranking network by accommodating heterogeneous measurement error.

In addition to its ability to accommodate non-sequential ranking structures, our three-step approach provides advantages not found in other existing mathematical approaches. In our first step, we enhance the initial estimate of the dominance potential, which is based only on direct pairwise interactions, by further incorporating information gained through the assumption of dominance transitivity. By converting from the estimated probabilities to odds and adjusting the odds based on interactions with third parties, we are able to accommodate apparent inconsistencies among the data, such as when member  $i$  usually wins its conflicts with member  $j$  and  $j$  usually wins its conflicts with member  $k$ , but  $k$  usually wins its conflicts with  $i$  (see [9]). In such cases, the estimated dominance potentials among the three individuals, after converting back from odds, will rightly tend to be closer to 0.5 than they would be had transitivity been neglected. Moreover, in our third step, we take into account the uncertainty in the data in order to provide a measure of precision for our rank estimates. An individual that consistently wins in many conflicts with individual  $i$  and consistently loses in many conflicts with member  $j$  should almost certainly be assigned a rank between the ranks of  $i$  and  $j$ . But for many individuals there are few observed conflicts with the other members, and the outcomes of those few conflicts may vary from one occasion to the next. In such cases the ranking coordinates are less certain. This uncertainty is quantified under our approach using repeated samples from a posterior probability distribution, so that confidence bounds for the ranks of individuals may be established. We are unaware of any other method that accounts for the uncertainty in this manner.

In an attempt to validate our proposed three-step approach and at the same time make a comparison with a result based on Bayesian analysis [6], we reanalyze a data set of twenty longhorn sheep from [5]. We show our steps are capable of computing a set of “perfect” non-sequential ranking networks, in which our estimate of  $P = [p_{ij}]$  contains no values in the lower triangle exceeding 0.5. The existence of such a set is primarily due to the fact that there are two sheep that have no direct interactions with most of the dominant sheep, but have a couple of wins over lower-ranking sheep. Therefore they can be placed at several different ranking network coordinates without affecting the non-sequential ranking network. This perspective of a ranking network is nearly impossible for a purely sequential ranking model based on Bayesian analysis, even for a relatively small society.

We apply our three-step approach to sets of conflict data, collected at CNPRC, by first establishing a ranking network for 94 adult rhesus macaques (mean total group size = 137 adults, juveniles, and infants) living in a half-acre outdoor captive enclosure. The resultant ranking network is coupled with known matriline information. We then apply our approach to conflict data collected from a second colony of macaques to compute the ranking network and confirm the utility of this methodology. In this fashion we address the question: Is the rhesus macaque’s society organized as a purely corporative ranking network, a kingdom version of a corporative network, or a mixture of these? The characterization of rhesus macaque society in terms of corporative or kingdom ranking networks may also shed light on how such hierarchical networks are maintained through lower-level interactions among individuals and groups of individuals (such as matriline).

## Methods

### Ethics statement

All research reported in this manuscript adhered to the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health, the laws of the United States government, and the recommendations of the Weatherall report, “The use of non-human primates in research.” All research subjects were housed in large social groups in half-acre outdoor enclosures to provide for their psychological well-being. The methodological approach was purely observational, and involved no experimental or invasive treatment of the animals. All occurrences of illness or injury among study subjects were immediately reported to and treated by CNPRC veterinary staff, and all efforts were made to ameliorate suffering. This project was approved by the University of California, Davis Institutional Animal Care and Use Committee, protocol #11843.

Consider a single enclosure with  $N$  adult rhesus macaques indexed by  $i$  from 1 through  $N$ . For any  $(i, j)$  pair of macaques, we want to infer the unknown probability  $p_{ij}$  that the  $i$ th macaque dominates the  $j$ th macaque, as well as the reverse dominance probability  $p_{ji} = 1 - p_{ij}$ . The inference for  $p_{ij}$  and  $p_{ji}$  is based on the behavioral records  $c_{ij}$  and  $c_{ji}$ , which count the numbers of observed wins for the  $i$ th macaque over the  $j$ th macaque, and vice versa, among their total number  $c_{ij} + c_{ji}$  of observed direct conflicts.

Due to the weekly behavioral sampling scheme (6 hrs/day, 4 days/week; see [11], in press, for a complete description of observational methods), it is not uncommon to find severe heterogeneity in the collection of data  $\{c_{ij} \mid i \neq j, i, j = 1, \dots, N\}$ . In a cage consisting of  $N = 94$  mature macaques, the total possible number of pairwise interactions within the social group is  $N(N - 1) = 94 * 93 = 8742$ . But only a small percentage of pairs were observed to have any direct interaction, and the remaining pairs were never observed to interact. Hence, under modeling with Bernoulli random variables, the unknown parameters  $\{p_{ij}\}$  would be estimated with very high heterogeneous precision for those cases of  $c_{ij} + c_{ji} > 0$ . However, their estimated dominance probabilities may differ significantly from 0.5, because of the transitivity property. Thus, the collection of  $N(N - 1)$  parameters are dependent upon one another in a non-sequential fashion, and not orthogonally, as is often assumed. The parameter space, where the “true” parameters live, is indeed a very complex manifold.

Given the non-sequential interdependence among the collection of parameters in such a network, the likelihood-based approaches, including maximum likelihood and Bayesian analysis, are inappropriate for proper inference on the  $N(N - 1)$  parameters, simply because the optimization task upon a nearly  $10^4$ -dimensional complex manifold is practically impossible. A glimpse of such difficulty can be seen in the Bayesian analysis reported in [6]. Even when ignoring the interdependence among the parameters, the high dimensionality alone, despite the small group size of  $N = 20$ , causes the posterior density to be too flat to be of practical use.

The first task in computing a non-sequential ranking network is to simultaneously calculate the  $N(N - 1)$  dominance potential parameters for all possible pairwise interactions among network members. Since many of these pairs have no observed direct interactions, the dominance potential parameters must be inferred from the transitivity relationships among the parameters. These features create very high dimensionality in the parameter space and result in high heterogeneity of precision in parameter estimation. We address this task in the first subsection.

After estimating the  $N(N - 1)$  parameters, the next challenging task is to construct a non-sequential ranking network. This task is analogous to working out a puzzle with many irregularly-shaped pieces and many missing pieces. We accomplish this in the second subsection by transforming the network construction task into an optimization problem, and then developing an approach equipped with Simulated Annealing [12] to resolve this fundamental construction. Finally, we present a third step in the third subsection in which we derive a basis for empirical confidence bounds for features observed in the computed ranking network.

**The first step: S1**

Given the  $N \times N$  matrix  $C = [c_{ij}]$  of behavior records for each pair of macaques, we construct four new  $N \times N$  matrices. Matrix  $W^0 = [w_{ij}^0]$ , with  $w_{ij}^0 = (c_{ij} + 1)/(c_{ji} + 1)$  for  $1 \leq i, j \leq n$ , gives the empirical odds that macaque  $i$  dominates macaque  $j$ , based on observed confrontations between the two. We add one to each record to avoid division by zero. Then matrix  $\hat{P} = [\hat{p}_{ij}]$ , where  $\hat{p}_{ij} = (c_{ij} + 1)/(c_{ij} + c_{ji} + 2)$ , gives the empirical probability that macaque  $i$  dominates macaque  $j$  based solely on direct observations of conflicts. The computed  $\hat{p}_{ij}$  serve as initial estimates of the unknown probabilities  $p_{ij}$ , which are the true pairwise dominance probabilities.

Here  $\hat{p}_{ij}$  may be regarded as the posterior mean of a beta distribution with parameters  $\alpha = c_{ij} + 1$  and  $\beta = c_{ji} + 1$ . That is, given a uniform prior distribution for the unknown probability  $p_{ij}$  and a binomial distribution for  $c_{ij}$  (with parameters  $c_{ij} + c_{ji}$  and  $p_{ij}$ ), the posterior distribution of  $p_{ij}$  given  $c_{ij}$  is the Beta  $(\alpha, \beta)$  with mean  $\alpha/(\alpha + \beta) = \hat{p}_{ij}$ . This idea will be exploited in step S3.

From  $\hat{P}$  we construct the transitivity structure matrix  $W^1 = [w_{ij}^1]$  such that

$$w_{ij}^1 = \prod_{h \neq i, j} \max\left(1, \frac{\hat{p}_{ih}\hat{p}_{hj}}{1 - \hat{p}_{ih}\hat{p}_{hj}}\right).$$

Here  $w_{ij}^1$  signifies the dominance odds between macaques  $i$  and  $j$  based solely on transitivity implied by their interactions with other single intermediate macaques. Note that  $w_{ij}^1$  is never less than one. In order for the ratio  $\hat{p}_{ih}\hat{p}_{hj}/(1 - \hat{p}_{ih}\hat{p}_{hj})$  to exceed one, the observed dominances of macaque  $i$  over intermediate macaque  $h$  and of macaque  $h$  over macaque  $j$  must both be strong enough to make the product  $\hat{p}_{ih}\hat{p}_{hj}$  exceed 0.5. Hence the dominance of macaque  $i$  over macaque  $j$  via transitivity is implied conservatively. The matrix  $W^1$  thus carries information about the dependence network structure embedded in the data.

Finally, we construct the matrix  $W^* = [w_{ij}^*]$  by first setting  $w_{ij} = w_{ij}^0 w_{ij}^1$  for  $1 \leq i, j \leq N$ , then setting  $w_{ij}^* = \sqrt{w_{ij}/w_{ji}}$  to standardize. This standardization makes  $w_{ij}^* = 1/w_{ji}^*$ . The matrix  $W^*$  thus gives the odds that macaque  $i$  dominates macaque  $j$ , implied by both observation of direct conflicts between them and by the transitivity relations among the other macaques. Step S1 outputs both  $W^*$  and  $W^1$  for use in the next two steps.

**The second step: S2**

The second step estimates the rank of each macaque in the social hierarchy based on the dominance odds computed in step S1. This is accomplished by first computing an initial ranking using the matrix  $W^*$  obtained in S1, then using the simulated annealing algorithm to improve upon this ranking by minimizing a cost function.

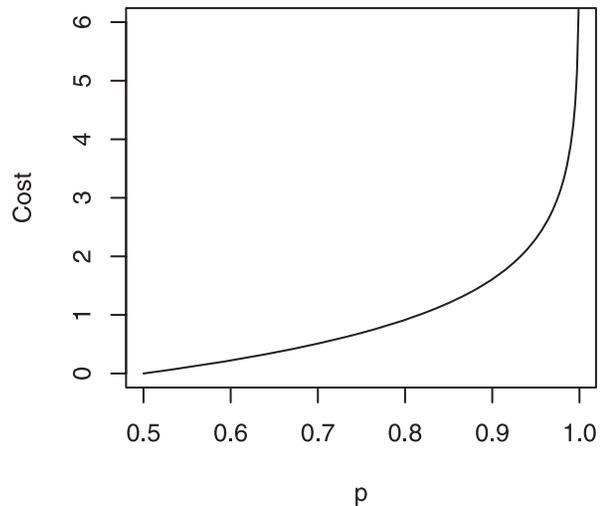
Based on the  $N \times N$  matrix  $W^*$  from S1, with elements  $w_{ij}^*$ , we construct the matrix  $P^* = [p_{ij}^*]$ , where  $p_{ij}^* = w_{ij}^*/(1 + w_{ij}^*)$  for

$i, j = 1, \dots, N$ . This converts the odds into enhanced probabilities. These probabilities  $p_{ij}^*$  differ from the probabilities  $\hat{p}_{ij}$  computed directly from the observed pairwise conflicts, since the dominance probability implied by transitivity has been incorporated into their computations. We then choose a threshold  $t$  and construct the matrix  $P^t = [p_{ij}^t]$  such that  $p_{ij}^t = 1$  if  $p_{ij}^* > t$  and equals zero otherwise. In practice we use  $t = 0.6$ . We seek to rearrange the columns of  $P^t$  such that the upper triangle contains few zeros and the lower triangle contains many zeros. To accomplish this, we compute the sum of each column, then order the columns such that the sums of the columns are nondecreasing from left to right. We then rearrange the rows and columns of  $P^*$  according to the same order. The purpose of this effort is to minimize the sum of those values in the lower triangle of  $P^*$  exceeding 0.5. The values in the lower triangle represent the estimated probabilities that lower-ranking macaques dominate higher-ranking macaques, which would not exceed 0.5 in the ideal situation. This technique produces the initial estimate of the macaque ranking order, which is not necessarily optimal.

To determine the quality of the estimated ranking order, we devise a cost function whose argument is the lower diagonal entries of the matrix  $P^*$ . We construct this function so that values greater than 0.5 in the lower diagonal will be penalized, with the penalty increasing at a greater rate as the values approach one. While other definitions are plausible, the cost function used here is defined as

$$C(P^*) = \sum_{i=2}^N \sum_{j=1}^{i-1} \max(0, -\log[2(1 - p_{ij}^*)]). \tag{1}$$

As Figure 2 shows, for each element  $p_{ij}^*$  in the lower triangle of  $P^*$ , this function adds cost equal to  $-\log[2(1 - p_{ij}^*)]$  whenever  $p_{ij}^* > 0.5$ , with the cost increasing steeply as  $p_{ij}^*$  approaches one. In the ideal situation, no element in the lower triangle of  $P^*$  exceeds 0.5, and thus the cost is zero. But observations of conflict behavior in a complex society may not allow the computation of an optimal cost near zero. The simulated annealing (SA) algorithm [12] is implemented to seek an optimal ranking order that will provide a total cost which is as close to zero as possible.



**Figure 2. Plot of cost function for  $0.5 < p < 1$ .**  
doi:10.1371/journal.pone.0017817.g002

In the SA algorithm, we start at an initial ranking order. We then consider a “neighbor” ranking order, in which the ranks of two randomly selected macaques are switched. If the cost function decreases when  $P^*$  is modified by rearranging the rows and columns according to the neighbor ranking order, we move to the new ranking order. Otherwise, we move to the new ranking order with probability  $\pi_k$ . We compute  $\pi_k$  based on our progress in the SA algorithm, so that if we plan to complete  $K$  iterations,  $\pi_k$  will decrease exponentially from nearly 1 to almost 0 as  $k$  approaches  $K$ . In practice we use  $K = 10,000$  and decrease  $\pi_k$  only when  $k$  is a multiple of 10.

After the last iteration we save the ranking order which produced the lowest cost encountered during the SA algorithm. The SA algorithm is designed so that it will tend to find the global minimum on a surface that may have multiple local minima. We run the SA algorithm many times and choose the best overall ranking order, i.e., that which corresponds to the rearrangement of rows and columns of  $P^*$  which yields the lowest cost. This order is taken as our estimate of the ranking of the mature rhesus macaques in a given enclosure.

**The third step: S3**

The third step is used to construct a basis for empirical confidence intervals for desired features of the estimated ranking order obtained from step S2. For all  $1 \leq j < i \leq N$  we randomly draw  $\tilde{p}_{ij}$  from the beta distribution with parameters  $\alpha = c_{ij} + 1$  and  $\beta = c_{ji} + 1$ . For  $j > i$  we set  $\tilde{p}_{ji} = 1 - \tilde{p}_{ij}$ . Note that the mean of the given beta distribution is  $\alpha / (\alpha + \beta) = (c_{ij} + 1) / (c_{ij} + c_{ji} + 2) = \hat{p}_{ij}$ . Hence  $\tilde{p}_{ij}$  represents a random perturbation of the empirical dominance probability  $\hat{p}_{ij}$  of macaque  $i$  over macaque  $j$ . Once we obtain the matrix  $\tilde{P} = [\tilde{p}_{ij}]$  we set  $\tilde{w}_{ij}^0 = \tilde{p}_{ij} / \tilde{p}_{ji}$  to convert the probability into odds, and thus construct the matrix  $\tilde{W}^0 = [\tilde{w}_{ij}^0]$ . Using the transitivity structure matrix  $W^1$  constructed in step S1,

which carries the transitivity structure that characterizes the real data, we construct the matrix  $\tilde{W}^* = [\tilde{w}_{ij}^*]$  by first setting  $\tilde{w}_{ij} = \tilde{w}_{ij}^0 w_{ij}^1$  for  $i, j = 1, \dots, N$ , then setting  $\tilde{w}_{ij}^* = \sqrt{\tilde{w}_{ij} / \tilde{w}_{ji}}$  to standardize. We run step S2 on  $\tilde{W}^*$  to obtain an optimal ranking order  $\sigma_1$ . We then repeat with another draw from the beta distribution to obtain  $\sigma_2$ , and so on to obtain  $B$  optimal ranking orders  $\sigma_1, \dots, \sigma_B$ .

We may then examine the distribution of the estimated rank for any individual macaque among the  $B$  ranking orders to derive confidence bounds for the rank estimate obtained in step S2. We can also derive bounds for the range of rankings corresponding to a specified matriline within the macaque society, or for other features of interest.

**Results**

**Female longhorn sheep**

A data set consisting of pairwise wins and losses within a group of  $N = 20$  female longhorn sheep originally from [5] was analyzed in [6] using Bayesian analysis under the sequential ranking network framework. The matrix in Table 1, reproduced from [6], shows in entry  $(i, j)$  the number of times sheep  $i$  defeated sheep  $j$  in confrontations. This matrix, arranged in ranking order, possesses several typical features of observational studies: few interactions for most pairs, and many pairs with no interactions. By assuming the existence of a sequential ranking hierarchy among the 20 female sheep, Bayesian analysis is conducted to estimate the twenty individual dominance potentials, say  $\{d_i\}$ ,  $i = 1, \dots, 20$ . The estimation is based on the likelihood constructed under an independence assumption among all pairs as

$$\prod_{i \neq j} K(d_i - d_j)^{c_{ij}}$$

**Table 1.** Conflict data for female bighorn sheep.

ID	15	17	21	06	08	01	10	22	02	11	05	16	12	09	26	25	29	27	28	24
15	0	0	4	0	7	0	0	1	1	0	3	0	4	1	1	2	2	9	2	5
17	0	0	3	0	1	0	0	0	0	1	2	0	5	1	0	0	0	2	1	1
21	0	0	0	0	2	0	1	1	1	1	1	1	1	1	3	0	1	3	2	1
06	0	0	0	0	0	0	0	1	0	0	0	0	3	0	1	0	0	1	1	1
08	0	0	0	1	0	1	0	2	3	2	1	0	1	1	2	2	2	6	4	0
01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
10	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0
22	0	0	0	0	0	0	0	0	2	1	2	0	1	0	2	0	0	1	2	3
02	0	0	0	0	0	0	0	0	0	0	1	2	1	4	1	0	0	1	0	0
11	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1
05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	1	1
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
12	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	1	1	1	0	0
09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Matrix of wins and losses for 20 female bighorn sheep, *Ovis canadensis*, with rows and columns arranged according to estimated dominance rankings [5,6]. doi:10.1371/journal.pone.0017817.t001

where the logistic version of the Bradley–Terry model [13] sets  $K(d_i - d_j) = 1/(1 + \exp[-(d_i - d_j)])$  as the kernel for transforming the difference between dominance potentials into a winning probability. The twenty-dimensional, mutually-independent normal flat priors of  $(d_1, \dots, d_{20})$  are used. An estimate of  $(d_1, \dots, d_{20})$  is permuted and gives rise to a sequential ranking hierarchy.

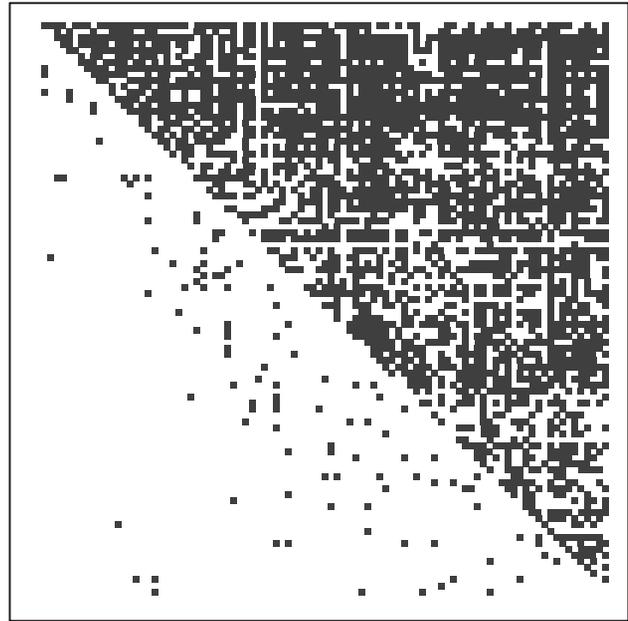
Unfortunately, the posterior density of  $(d_1, \dots, d_{20})$  from this Bayesian analysis is just too flat to give rise to a meaningful sequential ranking hierarchy. In contrast, our three-step approach computes several zero-cost permutations among the 20 female sheep. The most obvious one is (15, 17, 21, 08, 06, 01, 10, 22, 02, 11, 16, 12, 05, 09, 26, 25, 29, 27, 28, 24), which pushes down the locations of sheep numbers 06 and 05 in the Adams result. Indeed, when the data are displayed according to this permutation, all but one entry in the lower-triangle of the matrix of Table 1 are zeros. When the estimated dominance probability matrix  $P^*$  is arranged in this order, no values in its lower triangle exceed 0.5, so that its cost is zero. Hence this computational result implies that the above permutation produces a perfect ranking order, while the Bayesian result is flawed.

Typically a macaque conflict data set, as we will analyze in the next section, does not allow for such perfection when it comes to ranking. Usually there are many seemingly contradictory wins and losses that further complicate the task of ranking, particularly in species with frequent alliances and/or bidirectional aggression (see, e.g., [14]). However, such complexity is well-expected in a more sophisticated society. Hence it is deduced here that Bayesian analysis and the sequential ranking assumption are not suitable for a data set embedded with such kinds of complexity, especially when  $N$  is as large as 100.

### Ranking network for rhesus macaques

We apply our three steps to conflict data collected on two outdoor captive groups of rhesus macaques housed at the CNPRC. These behavioral data were collected between June 2008 and April 2009, and include all aggressive interactions that had a decisive outcome (for detailed methods, see [11]). For example, an interaction where an initiator threatens a recipient and the recipient runs away is counted as a win for the initiator and a loss for the recipient. Other aggressive behaviors include lunging, chasing, and biting. We apply steps S1 through S3 to these data. The first group (Cage 5) includes 94 adult rhesus macaques. Figure 3 is a graphical representation of the optimal dominance probability matrix  $P^*$  returned by S2. A black square indicates a matrix entry whose value exceeds the threshold of 0.6, which we take to be a strong indicator of dominance. The cost of the corresponding optimal ranking network, using relation (1), is computed to be 55.02. It is noted that the cost for the initial ranking order used in step S2 was 205.62. Thus we see that the Simulated Annealing algorithm achieves a very significant improvement. From the ranking perspective, this network reveals a clear non-sequential structure. Although it does not look exactly like the matrix shown in Figure 1(b), there are many areas in the upper triangle of Figure 3 that are rather sparse. Those sparse areas above the diagonal are likely to indicate parallel groups, while sparse triangular areas along the diagonal are rank-blurring groups. The latter groups' memberships are very essential for this non-sequential ranking network because they constitute the network's backbone structure.

Based upon the optimal ranking network, we list in Table 2 the rank coordinates computed for the members of each matriline in the macaque enclosure (Cage 5). Males are distinguished by bold italics. The most significant finding from this analysis is that the ranking network for this macaque society seems to have a



**Figure 3. Computed rhesus macaques' dominance probability matrix, with rows and columns arranged based on the optimal ranking order, for Cage 5.** A black square indicates a computed dominance probability above 0.6.  
doi:10.1371/journal.pone.0017817.g003

“governing” group consisting of a dominant matriline (C7) subordinate only to a lone alpha male, and several additional elite individuals of either gender from several lower-ranking matriline. This “power-sharing” structure seems to be very similar to human political systems. This sophisticated structure may be one of the key factors underlying the group's stability. Another significant finding, based on the sparse sections visible along the middle of the diagonal of Figure 3, is the existence of a subgroup of middle-ranking individuals who are somewhat parallel in dominance potential. These sparse sections correspond to the macaques between the 23rd and 50th ranks. Interestingly, this subgroup consists of a single middle-ranking matriline (F10), together with several individuals from other middle-ranking matriline (D10, S16, J6, L4), as may be concluded from Table 2. All these facts reveal the evident non-sequential structure in the ranking of the 94 individuals.

In terms of the corporate structure depicted in Figure 1(a), it appears that matriline C7 would correspond with group A at the top level, while matriline X1, S16 and J6 would correspond with groups B, C and D on the second tier. Matriline L10, D4 and F10 would share the third tier, corresponding to groups E, F and G. Matriline N4, M10 and Z2 would comprise a fourth tier that has no counterpart in Figure 1(a), while matriline G8 occupies the bottom level, corresponding to group H.

Based on our output from step S3, which involved 105 iterations, we may infer confidence bounds for the ranking of each macaque. Figure 4 displays boxplots for the top 20 macaques based on the optimal ranking determined by step S2 (indicated by open circles in the figure). Each boxplot indicates the distribution of computed rankings for each of these macaques over the 105 iterations. Points above or below the whiskers of a boxplot are outliers. A short box suggests that the corresponding macaque's rank is fairly consistent over the 105 iterations, as is the case with the alpha male (ID 24926). Hence the rank of such individuals is

**Table 2.** Cage 5 rankings by matriline.

Matriline	Ranking coordinates
NRM	<b>1</b>
C7	3 4 <b>5</b> 6 7 8 9 10 12 14 <b>15</b>
X1	11 16 17 20 24 69 <b>89 92 94</b>
S16	18 <b>19</b> 21 22 23 <b>35 42</b>
D10	26 <b>34</b>
J6	25 29 31 36
L4	30 33 <b>38</b>
F10	27 28 <b>37</b> 39 <b>40</b> 41 43 44 45 46 47 <b>48 49</b> 50 51 52 <b>84</b>
N4	32 <b>53</b> 54 55 56 57 58 59 60 <b>61</b> 63 <b>65</b> 68
M10	<b>2</b> 62 64 <b>66 76 77</b>
Z2	<b>13</b> 67 70 71 72 73 74 78 <b>79 80</b> 81 82 83 85 86
G8	<b>75</b> 87 88 90 91 93

Summary of ranking coordinates for rhesus macaques in Cage 5 grouped according to matriline. Males are indicated by bold italics. doi:10.1371/journal.pone.0017817.t002

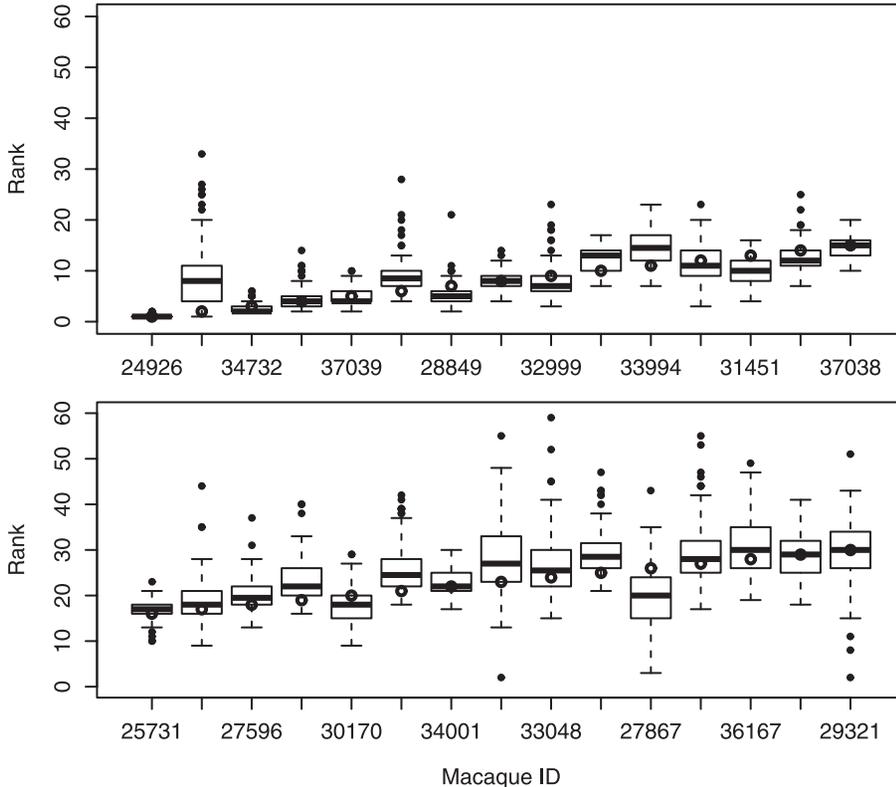
fairly certain. A long box suggests that the corresponding macaque’s estimated rank is quite uncertain, given the data, as with the beta male (ID 22898). Moving this macaque’s position in the ranking order does not affect the cost significantly. Since there are few such individuals in this example, the variability of their rankings over the 105 iterations has only a slight impact on the variability of the rankings for their neighbors in the hierarchy.

We might also take interest in the range of ranks for a group of macaques, such as for a particular matriline. Our step S2 estimated the ranks for the eleven macaques in the C7 matriline to fall between 3 and 15 inclusive, making this the highest-ranking matriline in the society. Using the output of step S3, we find that the highest rank assigned among the members of C7 was distributed as follows: 1 (one time), 2 (88 times) and 3 (16 times). Meanwhile, the lowest rank among the C7 macaques was distributed more widely:

Rank	13	14	15	16	17	18	19	20	21	22	23	25	28
Frequency	3	13	27	23	12	12	4	4	2	1	2	1	1

The range between the highest and lowest ranks was centered at 14. To obtain empirical 95% confidence bounds, we eliminate the 5 most extreme values from each distribution, and conclude with 95% confidence that the highest rank for matriline C7 is between 2 and 3, while its lowest rank is between 13 and 22. This gives strong evidence that this matriline dominates the others.

We implement the same procedure for a second macaque colony living in a separate enclosure (Cage 8) at the CNPRC. We initially identify 136 adults in the colony. However, only 91 of these adults were involved in observed conflicts during the period of interest, so the other 45 adults were removed from consideration. The value of the cost function (1) for the optimal ranking network for these macaques, based on the output of step S2, was 49.59. Once again, a lone male is identified at the top of the hierarchy, followed by a dominant matriline (C16). The remaining matriline follow in succession, with several matriline having an elite member of either gender whose rank is very high



**Figure 4.** Boxplots indicating confidence bounds for top twenty rhesus macaques, based on optimal ranking order computed in step S2 and 105 iterations of step S3 for macaques in Cage 5. doi:10.1371/journal.pone.0017817.g004

**Table 3.** Cage 8 rankings by matriline.

Matriline	Ranking coordinates
S22	<b>1</b>
C16	2 3 6 7 8 9 10 <b>13 14</b>
F13	4 11 15 16 17 18 19 20 21 22 23 <b>43 87</b>
V5	24 25 26 27 28 29 30 31 34 35 <b>90</b>
L8	36 37 38
K8	12 39 40 41 42 44 45 46 47 49 52 69 <b>86 91</b>
D13	<b>5 48 54</b>
G11	51 55 58 <b>59 60 61 64 89</b>
Q3	50 53 56 57 63 85 <b>88</b>
T9	65 66
B12	67 68 70 71 72 <b>81</b>
M14	32 73
N8	<b>62 74 75 77 78 79</b>
N6	<b>33 76 80</b>
P11	82 83 84

Summary of ranking coordinates for rhesus macaques in Cage 8 grouped according to matriline. Males are indicated by bold italics.  
doi:10.1371/journal.pone.0017817.t003

compared to the rest of the matriline. The ranks are arranged by matriline in Table 3, with males indicated by bold italics. In terms of Figure 1(a), we do not find quite the same multi-tiered structure. It is likely matriline G11, Q3 and D13 occupy the same tier, as do matriline N8 and N6, while the other matriline occupy tiers by themselves. Since the analysis of the second colony identifies a similar network structure, our confidence in the consistency of this three-step approach is strongly validated. Our analysis of several additional colonies yield comparable results.

## Discussion

We propose a three-step approach to compute and construct a non-sequential ranking network based on highly heterogeneous counts of aggressive interactions among mature members of a society of rhesus macaques. We provide theoretical arguments and explicit reasoning that describe the challenging nature of this task. We propose a set of three steps, each of which is designed to resolve a significant challenge encountered in the construction of real-world ranking networks. However, our solution to this computational challenge sheds light on how the underlying corporative kingdom structure of rhesus macaque society influences the behavior of individuals as well as the stability of the society as a whole. We expect that this new approach will be useful in many other studies, especially when dealing with the property of transitivity among binary relational data. From this perspective, corporative network theory may still be in its infancy.

Finally, we briefly comment on the limits of the maximum likelihood estimate (MLE) approach and Bayesian analysis in constructing corporative ranking networks. The transitivity-

induced parameter dependence makes the pertinent parameter space a complex manifold. This manifold structure is problematic for the optimization procedures of the MLE and the assignment of reliable prior distributions in Bayesian analysis. In particular, Bayesian analysis always produces a very flat posterior density when employing a non-informative prior on even a modest number of parameters. In addition, the high dimensionality of this manifold structure further complicates the situation, if not rendering it impossible to resolve. For example, it is common practice among animal behaviorists to impose a sequential ranking assumption on the structure of animal societies [6]. However, the imposition of a sequential rank order ignores the aforementioned difficulties. Indeed, an assumption of a sequential structure implies transitivity, which further necessitates dependence among the parameters, resulting in such difficulties. Therefore, it is our hope that the three steps developed here will be broadly applicable to the study of ranking network structures of many societies.

This three-step approach is an advancement in the challenging task of computing rank coordinates in societies which do not possess a sequential hierarchy. By incorporating information gained through dominance transitivity, we accommodate apparently contradictory cycles in the data to obtain reasonable estimates of pairwise dominance probabilities. We also acknowledge the uncertainty in the data and exploit it so as to provide reasonable confidence bounds for the ranks of individuals. While other mathematical methods assume a sequential ranking order for all members of an animal society, our approach provides more flexibility, in that the computed result allows for some variability in the estimated ranking coordinates. This is because the assigned ranks for some individuals may be interchanged without affecting the value of the cost function.

For societies like the rhesus macaques, this flexibility in our approach is essential. The scientists who observe the macaque societies at the CNPRC are in agreement that their hierarchies are not sequential. The relative ranks of matriline can be identified by observation, but the ranks of individuals within the matriline are more difficult to identify, partly because the groups are very large, and partly because some matriline are genetically fragmented. Hence, while the imposition of an assumption that the hierarchy is sequential is necessary under other accepted ranking methods, it is artificial, and the output from such methods generally conflicts with the observations of researchers. A method based on a corporative kingdom model, as presented here, is coherent with these observations, and is therefore a more realistic tool for the analysis of rhesus macaque ranking networks.

## Acknowledgments

The authors gratefully acknowledge the contributions of Megan Jackson, Shannon Seil, Ashley Cameron, Allison Heagerty and Christopher Dienes.

## Author Contributions

Conceived and designed the experiments: BM BB HF MPA. Performed the experiments: BM BB HF MPA. Analyzed the data: HF MPA. Contributed reagents/materials/analysis tools: BM BB HF MPA. Wrote the manuscript: HF MPA BB BM.

## References

- Drews C (1992) The concept and definition of dominance in animal behavior. *Behaviour* 125: 283–313.
- Bonabeau E, Theraulaz Z, Deneubourg JL (1996) Mathematical model of self-organizing hierarchies in animal societies. *Bulletin of Mathematical Biology* 58: 661–717.
- Goessmann C, Hemelrijk C, Huber R (2000) The formation and maintenance of crayfish hierarchies: behavioral and self-structuring properties. *Behavioral Ecology and Sociobiology* 48: 418–428.
- Nagy M, Ákos Z, Biro D, Vicsek T (2010) Hierarchical group dynamics in pigeon flocks. *Nature* 464: 890–893.

5. Hass C (1991) Social status in female longhorn sheep (*Ovis canadensis*): expression, development and reproductive correlates. *J of Zoology* 225: 509–523.
6. Adams E (2006) Bayesian analysis of linear dominance hierarchies. *Animal Behaviour* 69: 1191–1201.
7. Isbell L (2002) Ecological models of female social relationships in primates: similarities, disparities and some directions for future clarity. *Behaviour* 139: 177–202.
8. Hemelrijk C, Wantia J, Gygax L (2005) The construction of dominance order: comparing performance of five methods using an individual-based model. *Behaviour* 142: 1037–1058.
9. De Vries H (1998) Finding a dominance order most consistent with a linear hierarchy: a new procedure and review. *Animal Behaviour* 55: 827–843.
10. Wasserman S, Faust K (1994) *Social Network Analysis*. New York: Cambridge University Press.
11. Beisner B, Jackson M, Cameron A, McCowan B (2010) Effects of natal male alliances on aggression and power dynamics in rhesus macaques. *American Journal of Primatology*, In press.
12. Kirkpatrick S, Gelatt C, Vecchi M (1983) Optimization by Simulated Annealing. *Science, New Series* 220: 671–680.
13. Bradley R, Terry M (1952) Rank analysis of incomplete block designs. I. The method of paired comparison. *Biometrika* 39: 324–345.
14. Thierry B, Singh M, Kaumanns W, eds. *Social epigenesis*. Cambridge: Cambridge University Press.