ANIMAL BEHAVIOUR, 2008, 75, 1809–1815
doi:10.1016/j.anbehav.2007.10.029

Available online at www.sciencedirect.com
ScienceDirect

COMMENTARY

Incorporating uncertainty into the study of animal social networks

DAVID LUSSEAU, HAL WHITEHEAD & SHANE GERO
Department of Biology, Dalhousie University

(Received 26 January 2007; initial acceptance 7 April 2007; final acceptance 8 October 2007; published online 14 March 2008; MS. number: AS-1311R2)

Keywords: bootstrap; social behaviour; social network; social structure

Network analysis is rapidly establishing itself as a powerful tool for studying the structure and dynamics of complex systems (Albert & Barabasi 2002; Newman 2003). It has proven useful in understanding social interactions among humans and nonhumans and how global properties emerge from them (Watts et al. 2002; Lusseau & Newman 2004; Flack et al. 2006). It has also been helpful in describing and predicting the behaviour of technological networks and some biological systems for which all interactions can be described as known absolute values. However, the application of network analysis to social systems involving nonhuman organisms has been slower, because it has been difficult to infer the statistical and biological significance of observed network statistics and structures (Croft et al. 2005; Lusseau et al. 2006).

Two key aspects have presented difficulties. First, in contrast to some human studies, analysts estimate social relationships among individuals, they do not know them, and often they estimate those based on quite limited data. Researchers estimate relationships by observing interactions or associations between individuals, ranging from behavioural events (such as grooming) to co-occurrence. They can then build relationship measures using interaction rates or association indexes (Whitehead & Dufault 1999). Yet these observations do not represent all the interactions occurring between individuals; they are a sample. Studies in animal network analyses have never discussed sampling uncertainty even though its consequences can greatly affect the results of such analyses when sample size (i.e. the number of times that individuals are observed) is small. For example, if two individuals are together 50% of the time, they have a true association index (Cairns & Schwager 1987) of 0.5. If they were identified together 10 times, the 95% confidence interval for the estimated association index is about 0.3–0.7 (Whitehead 2008).

A second problem is that most network analyses of nonhumans have focused on binary networks, in which relationships are defined as being either present or absent. The matrix that represents the network contains only ones (when two individuals are defined as associated) and zeros (when they are not). Researchers have used binary transformations of continuous matrices of interaction rates or association indexes to describe animal social networks. These transformations require certain arbitrary manipulations that can be justified to varying degrees (Lusseau 2003; Croft et al. 2005). For example, one might decide that association indexes smaller than an arbitrary value (say 0.5) should indicate the lack of a relationship between individuals (assigned a value of zero in the binary matrix) and those greater than 0.5 should indicate a relationship between individuals (assigned a value of one in the binary matrix). Another example is to define pairs for which the association index is greater than expected if interactions occurred by chance as relationships (ones) and others not possessing relationships (zeros). Authors largely ignore these manipulations when considering the conclusions derived from the results of these studies. In addition, most of these animal social systems are densely connected, and discarding information about the strength of relationships might significantly distort the interpretation of the network topology. In many nonhuman communities, all individuals associate with all other individuals at some rate, so with complete sampling and association used to indicate relationships, the binary network would link all individuals to all others. Different

Correspondence and present address: D. Lusseau, University of Aberdeen, Institute of Biological and Environmental sciences, Tillydrone Avenue, Aberdeen AB24 2TZ, U.K. (email: d.lusseau@abdn.ac.uk). Hal Whitehead and Shane Gero are at the Department of Biology, Dalhousie University, 1355 Oxford Street, Halifax, NS B3H 4J1, Canada.

0003–3472/08/$34.00/0

sampling rates and criteria for judging a dyad as linked can greatly change the perceived structure of a network (Croft et al. 2005). Binary simplification can lead to wrong interpretations about the social structure of the population or inappropriate divisions when defining community structure. It can also lead to wrong inferences about the position of individuals within the network.

We can also define networks with links between individuals representing the weight of associations between those individuals. These weighted networks can represent the matrix resulting from observations of interactions between or associations among individuals in the wild. Recent advances in weighted network analyses provide new tools to quantify the position of individuals in weighted networks and the community structure of those networks (Barrat et al. 2004; Newman 2004a, 2006b). In our view, these tools are particularly appropriate for the analysis of nonhuman social networks. However, a shift towards weighted networks in animal behaviour requires tools to deal with sampling issues. Here we introduce bootstrapping techniques to incorporate sampling uncertainty when estimating weighted network measures. We also introduce techniques that randomize networks subject to constraints to assess how data structure influences the observed statistical properties of networks. We use two examples to illustrate the value of these new techniques. First, we determine variation in network centrality measures between individuals within a small sperm whale, *Physeter macrocephalus*, social unit. We then apply these methods to assess the uncertainty surrounding community structure in a bottlenose dolphin (*Tursiops* sp.) population residing in Doubtful Sound, New Zealand (Lusseau 2003). Finally, using this bottlenose dolphin social network, we test how transitivity in association departs from random. These analyses were implemented in Matlab using the Socprog package, which is freely available at http://myweb.dal.ca/~hwhitehe/social.htm (Whitehead 2006).

Defining Weighted Networks

Nonhuman societies, ranging from social insects to mammals, are commonly studied using dyadic association data; that is, observations of interactions between pairs of individuals (Whitehead 1997; Whitehead & Dufault 1999; McComb et al. 2000; Watts 2000; Shimooka 2003; Sigurjonsdottir et al. 2003; Boogert et al. 2006; Greene & Gordon 2007; McDonald 2007). Association measures should indicate whether a pair of animals is in circumstances in which they may behaviourally interact (Whitehead & Dufault 1999), and these measures are often based upon common membership of transitive groups or other symmetric measures (e.g. within *x* body lengths), but asymmetric association measures are possible (e.g. nearest neighbours). We limit our explanation to the former type of data because the analysis of asymmetric association data requires further manipulations of network statistics that are beyond the scope of this study.

Analysts record associations among animals in sampling periods then use these data to calculate association indexes (Cairns & Schwager 1987), which vary from 0 (never found associated) to 1 (always found associated). The resulting association matrix is the basis of many traditional analyses of nonhuman social structures (Pepper et al. 1999; Whitehead & Dufault 1999) and it also defines a weighted network. In a display of this network, nodes represent individuals and linking edges have line widths proportional to the association index between the two individuals (e.g. Fig. 1).

Incorporating Uncertainty in Centrality Measures

The patterns of interactions within small social communities are difficult to quantify because of the issues associated with statistical inference based on a small number of data points (individuals in this case). It can therefore be difficult to understand whether different individuals play different structural roles within these units (Lusseau 2007a). Sperm whales live in matrilineal populations and females spend most of their lives within their natal units (Whitehead 2003). However, the structure of social relationships within these social units is not clear (Christal & Whitehead 2001). Matrilineal social units in sperm whales function to provide care for calves at the surface while mothers make deep dives for food (Whitehead 2003). As such, a calf should be a central focus of the unit's underlying social relationships to maximize the likelihood that it will survive.

We examined this issue using data collected on a social unit, the Group-of-Seven (GOS), in an area that covered approximately 2000 km^2 of the Commonwealth of Dominica (Gero 2005). The GOS consists of five adult females, one juvenile male (no. 5727, 8–10 years old) and one male calf (no. 5703) whose mother was no. 5722. Following previous studies (Whitehead 2003), we considered that individuals photo-identified together in clusters, defined as individuals that were within approximately three adult body lengths from any other member and coordinated in their behaviour. We used a half-weight

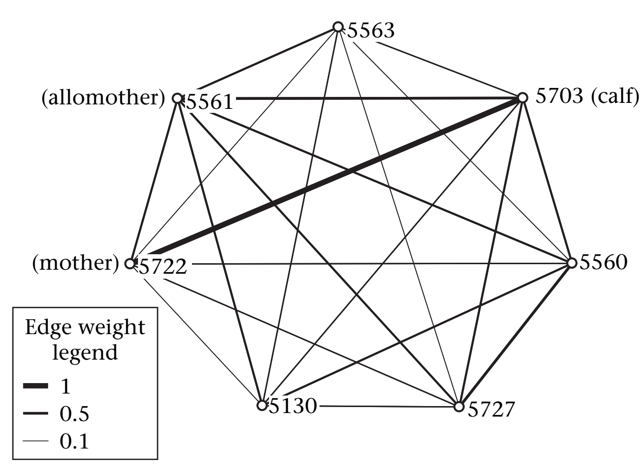

Figure 1. The association social network of the Group-of-Seven social unit of sperm whales. The thickness of the lines (edges) represents the weight of the association index (half-weight index).