Social Networking

Recent Trends, Emerging Issues and Future Outlook

Xin Ming Tu • Ann Marie White • Naiji Lu
Editors

MEDIA AND COMMUNICATIONS
Technologies, Policies and Challenges

NOVA

MEDIA AND COMMUNICATIONS - TECHNOLOGIES, POLICIES AND CHALLENGES

SOCIAL NETWORKING

RECENT TRENDS, EMERGING ISSUES AND FUTURE OUTLOOK

XIN MING TU
ANN MARIE WHITE
AND
NAIJI LU
EDITORS

nova publishers
New York

Library of Congress Cataloging-in-Publication Data

Social networking : recent trends, emerging issues and future outlook / editors, Xin Ming Tu, Ann Marie White and Naiji Lu.
pages cm
Includes bibliographical references and index.
ISBN: 978-1-62808-529-7 (hardcover)
1. Social networks--Research. 2. Online social networks--Research. I. Tu, Xin M. II. White, Ann Marie. III. Lu, Naiji, Ph.D.
HM741.S6346 2013
302.30285--dc23

2013026853

Published by Nova Science Publishers, Inc. †New York

MEDIA AND COMMUNICATIONS - TECHNOLOGIES, POLICIES AND CHALLENGES

SOCIAL NETWORKING

RECENT TRENDS, EMERGING ISSUES AND FUTURE OUTLOOK

MEDIA AND COMMUNICATIONS - TECHNOLOGIES, POLICIES AND CHALLENGES

Additional books in this series can be found on Nova's website under the Series tab.

Additional e-books in this series can be found on Nova's website under the e-books tab.

COMPUTER NETWORKS

Additional books in this series can be found on Nova's website under the Series tab.

Additional e-books in this series can be found on Nova's website under the e-books tab.

PREFACE

"Social networks" is no longer a term solely of the academe. A Google search of the term "social networks" at the time of writing this Preface yielded over 117,000,000 hits. Searching for "network analysis" yielded 6,460,000 hits. This is a field of study drawing upon graph theory to represent relationships between objects (e.g., websites) and how these form functional systems. "Social network analysis," the counterpart of network science focused on human interactions, yielded 939,000 or close to one million hits. Arguably, the advent of social media created a cultural meme that carries network concepts forward to general publics.

Social network analysis is a robust means to measure and map interactive systems. An approach to dynamically represent interactions among people and how these are organized – this method can capture ties within a system (e.g., between people) and between systems (e.g., between humans and other species or between humans and their environments). Due to inherent flexibility, network analytic approaches are now ubiquitous across any of a number of theoretical and empirical efforts in fields too numerous to list.

Social network analysis (SNA) focuses on social relationships (e.g., friendship) diagramed as nodes (points) and links (ties or edges between points). SNA examines features or changes to a social system illuminated from interactions and how these change over time (e.g., spread of disease). Social network analysis can measure and map "connectedness" or "flows" (e.g., information, resources, etc.) both within and across individuals, groups and organizations, and can locate these in a virtual- or geo-temporal space – yielding rich mixed methods possibilities (e.g., tying together joint analysis of physical or place-based attributes with social network attributes).

A notion that originated in sociology over a century ago, a social network is, at its heart, a social science theoretical concept that has widely spread to become an interdisciplinary application. As a method, it has yielded highly influential studies that generate implications beyond qualitative descriptions. For instance, Christakis and Fowler's 2007 article in the *New England Journal of Medicine* that drew upon Framingham Heart Study data, helped mobilize obesity as a public health priority. While in essence caused by behaviors not pathogens, the use of network analyses helped characterize obesity as contagious, or a spreading disease, and effectively conveyed the notion of an 'obesity epidemic' to rally greater public concerns.

Relationships between actors in a social group comprise a general mechanism underpinning any of a number of key accumulations or traits attributed to a social context or system – such as capital, efficiency or optimization, evolution, ecology, and spreading of information or ideas. As the referent of "social networks" (SN) has spread in academic fields as well as general everyday use, so have methods for illuminating and testing its properties and applications. From the early sociogram methods in sociology, to the advent of sophisticated visualization techniques, and most recently to its application in big data mining, the pace of innovation in SN methods and applications of concepts continues unfettered.

The growth and application of methods for studying social networks continues to burgeon in many fields. Given the scientific advancements in the study of networks over recent decades, highlights of this approach are now easily found among, and generated by, business, health, environment, computer science, statistics, economics and biology applications. Together, social network concepts and related analytic methods have propelled many fields towards greater understanding of critical phenomena such as infection, popularity or influence (e.g., page rank), and efficiency or capacity (e.g., transportation or trade).

However, applications of methodological advancements in many fields lag behind the pace of innovation. And fundamental limitations to surmount, to give this method greater explanatory (as compared to descriptive) power, have remained. Thus, this book is designed to promote wider reading about social network advancements across fields to accelerate the pace of interdisciplinary explorations and potential new discoveries.

The extensiveness and spread of SN theory and SNA's use and application suggests that emerging work in the many disciplines applying this field can be culled and encapsulated to advance research and training widely. Cross-pollination with other fields further astride in SNA is needed as many sciences (e.g., child development) have omitted foundational concepts in basic preparation of their future scientists. Our main objective for this volume is thus to share a wide set of field-specific insights that have the potential to advance other fields embracing similar foci.

We do so by first discussing conceptual issues in SNA, including the potential for applications of SN concepts yet to be addressed or encapsulated for wide dissemination within a particular field of study. For instance, a chapter in this volume poses answers to the question of what form SNA takes across the "observational-experimental" research design "continuum," an important topic in public health sciences. Basic questions about the power or utility of SN concepts are advanced via literature reviews (e.g., the role of opinion leaders in purchase behaviors).

Our volume also turns to methodological applications and advances in SNA. Currently, computer scientists increasingly seek to bring statistical expertise to bear in network analysis due to a host of unknowns to surmount in SNA. For instance, how do different strategies to identify a group affect conclusions about membership? How can we address a forgotten, but omnipresent, dimension of network datasets for valid statistical and causal inference (violation of independence assumption due to shared group membership) with SN datasets? Explorations to generate potential solutions or applications are presented.

Our goal for this book is to capture – across a wide range of fields – how emerging issues in the application of SN theory and SNA are being addressed. How to press forward past edges of our knowledge are illuminated as such a diverse set of authors' disciplinary expertise are brought to SNA. Each chapter selected illuminates new trends or applications that may have wide potential impact in other disciplines. Conceptual advances (e.g., applying the concepts of social networks such as peer influence on disease management and treatment adherence) and novel analytic approaches for studying properties of social networks are both highlighted.

These chapters convey that many frontiers remain for the study and application of social networks. Similar to a Google search described above, our call for papers shows that academic fields are at varying different stages of application of SNA. In our preparations for this book, we thus sought a call for papers that in fact yielded great diversity. We decided early on, that it is important for any book on SNA to share insights widely across fields.

Taken together, a picture emerges from the selected chapters that much work remains, work that can be cross-fertilized through multidisciplinary team building. The potential for this kind of team assembly grows as SN and SNA ideas spread and find fertile ground within respective fields. This volume demonstrates that with such a base, the next generation of SN concepts and methods can be propelled if the coalescing of interdisciplinary teams is fostered.

On a practical note, our efforts would not be in vain, if any reader "captures" an idea for use in their thinking or practice, or even if a reader seeks to connect to the work and commitment of these authors advancing SNA in each of their fields. Towards these ends – we invite readers to contact us as this volume "spreads" or influences yours or others attitudes, skills, behaviors, relationships and actions. We post our email contact information below. We look forward to hearing from readers about your use of this volume and working with you to grow this network of science.

Sincerely,

Naiji Lu
Department of Biostatistics and Computational Biology
Naiji_Lu@urmc.rochester.edu

Ann Marie White
AnnMarie_White@urmc.rochester.edu

Xin Tu
Xin_Tu@urmc.rochester.edu
Department of Biostatistics and Computational Biology
Rochester New York
University of Rochester

May 2013

CONTENTS

In: Social Networking
Editors: X. M. Tu, A. M. White and N. Lu
ISBN: 978-1-62808-529-7

Chapter 1

THE EFFECT OF FILTERING ON ANIMAL NETWORKS

Nienke Alberts, Stuart Semple and Julia Lehmann
Centre for Research in Evolutionary and Environmental Anthropology,
University of Roehampton, London

ABSTRACT

The past few years have seen a surge in the use of social network analysis to study animal sociality. Because ties between animals are usually inferred from behavioural interactions or spatial proximity, animal social networks may contain ties that are due to chance rather than representing a true 'bond'. To help focus on relationships that are more likely to be biologically meaningful, networks are often filtered by removing all the ties that are under a certain cut-off value. Researchers have proposed various methods to determine the level of filtering; however, it is not clear how these different methods of filtering may alter network metrics and consequently how they may affect the conclusions that are drawn from subsequent analyses. We investigate the effect that five commonly used filtering methods have on standard network metrics. To this end, social networks were generated using association indices of a troop of wild olive baboons (*Papio anubis*). These networks were filtered (i) until the network structure was significantly different from random, (ii) by median association strength, (iii) by mean association strength, (iv) until the giant component was close to breaking up, (v) by including only preferential associations. Global network metrics, individual network positions, and the extent of substructuring were determined and compared across the five filtered networks and the unfiltered network. Our results show that while global network metrics and individual network positions are affected by different filtering methods in a relatively predictable way, the number of substructures that were found in networks was strongly influenced by the way filtering was done. These results show that the appropriate filtering method needs to be carefully considered, based on the nature of the biological questions being asked.

INTRODUCTION

The last decade has seen a marked increase in the use of social network analysis in the *study of animal societies (Krause et al., 2009; Brent et al., 2011) across a variety of taxa*, ranging from fish (e.g. Croft et al., *2004*; Croft et al., *2006) to primates (e.g. Sueur and Petit* 2008; Lehmann and Boesch 2009; Ramos-Fernández et al., 2009; Henkel et al., 2010; Lehmann and Ross 2011), from elephants (e.g. Wittemyer et al., 2005) to squirrels (e.g. Manno 2008). This approach has proven valuable in addressing a wide range of questions about animal societies. For example, social network analysis has been used to characterise the social structures and identify substructures in animals groups in order to make fine-grained comparisons between populations, species and across time, (Lusseau et al., 2006; Sundaresan et al., 2007; Kasper and Voelkl 2009) and to assess the effect of environmental factors on social structure (Wittemyer et al., 2005; Henzi et al., 2009; Alberts 2012). The social network approach is frequently used to study animal populations with high levels of fission-fusion dynamics, *i.e.* where groups frequently split and reform, and where social structure is therefore not readily apparent (Ramos-Fernández et al., 2006; Sundaresan et al., 2007; Wolf and Trillmich 2008; Ramos-Fernández et al., 2009). In such populations, social network analysis is often used to identify layers in the social structure that were hitherto unknown (Lusseau et al., 2006; Ramos-Fernández et al., 2006; Sundaresan et al., 2007; Wolf et al., 2007). The social network approach has also been used for the identification of stable social bonds between individuals, and investigation of how these bonds may benefit individuals (Croft et al., 2004; Lehmann and Boesch 2009; Lea et al., 2010; Brent et al., 2011). An important use of social network analysis is to investigate the role of individuals within networks (Mitani 1986; Flack et al., 2006; Sueur and Petit 2008; Ramos-Fernández et al., 2009; Henkel et al., 2010), and in particular to explore how individual characteristics, such as sex or age, may influence an individual's position in their social network (Blumstein et al., 2009; Ramos-Fernández et al., 2009; Lehmann and Ross 2011).

An important issue that has arisen from studies of animal social networks in the last decade is how relationships between individuals are defined, and what constitutes a 'tie'. In animal social networks, relationships between individuals are inferred, usually from behavioural interactions or from spatial proximity. For some behavioural interactions, such as grooming, inferring social relationships can be relatively straightforward, as individuals target particular group members, and may make a considerable investment of time in such an interaction. However, inferring relationships from other behaviours may result in social networks that contain ties that do not represent a true 'relationship', but rather are due to chance encounters. In particular, networks that are based on associations defined by 'the gambit of the group' (Whitehead and Dufault 1999), *i.e.* when individuals are assumed to be associated when they are found in the same group, are more likely to contain ties that are due to chance events (Croft et al., 2008). In addition, unlike human groups, in many taxa societies are relatively small and closed, so that there is a clearly demarcated group of individuals that belong to the network, and changes to the composition of the network only occur through demographic changes (*i.e.* births, deaths, emigrations, immigrations). Such societies may lead to social networks in which all, or the majority of individuals are interconnected (Jacobs and Petit 2011), making it impossible to differentiate between network metrics of different networks if weighted networks are not used. In such weighted metrics, the strengths of the relationships are indicated by the weight of ties. The development of network metrics that specifically take into account the weights of ties, such as those implemented in *tnet* (Opsahl 2009), has been an important advance in this field. Such weighted network metrics can, for example, help to differentiate between an animal that is very social, and thus has many strong ties, and an animal that has many weak relationships. Nevertheless, the majority of network metrics currently do not take into account the weights of ties. Thus, even when metrics are calculated in a weighted network, network metrics often only take into account whether ties are absent or present, and not the weights of ties.

To help differentiate between chance ties and real relationships, in other words to help focus on relationships that are more likely to be biologically meaningful, animal social networks are often filtered until non-random core elements remain (Croft et al., 2008). In these cases, a filter is applied to a network by removing all the ties that are under a certain cut-off value, which can in principle be set at any level (Croft et al., 2008; Sueur et al., 2011). While it is common to filter animal networks prior to analysis, there is no set method for determining the appropriate cut-off value that should be used (Croft et al., 2008), and several different approaches have been proposed. Methods include filtering the network until its structure, which is measured by certain network metrics as test statistics, is significantly different from a random network structure (Brent 2009; Alberts 2012), filtering by median or mean association strength (Croft et al., 2004; Croft et al., 2008), or filtering until the giant component of the network is close to breaking up into smaller components (Croft et al., 2008). Perhaps the most frequently used method for the filtering of animal networks is to include only preferential associations (Whitehead 1999; Lusseau 2003; Wittemyer et al., 2005; Lusseau et al., 2006; Williams and Lusseau 2006; Sundaresan et al., 2007; Manno 2008; Lehmann and Boesch 2009; Ramos-Fernández et al., 2009; Henkel et al., 2010). In those studies, an association is considered 'preferential' when individuals associate significantly more frequently than predicted if individuals associated with each other at random. The choice of filtering methods is usually based on the type of data, *i.e.* binary or weighted, used in the study as well as the research questions that are addressed. While filtering of networks may be appropriate in studies that focus on the social relationships of individuals, it may not be for studies that focus on the transmission of disease or parasites (e.g. Corner et al., 2003; Cross et al., 2004; Godfrey et al., 2009) in which rare chance events may be very important. Researchers of animal social networks frequently use global network metrics to indicate the structure or qualities of the network to draw conclusions about their study groups. Additionally, researchers often determine individual network positions, and how these can be predicted by individual characteristics. Finally, social network analysis has been popular in helping to determine the level of substructuring of the network. Currently it is not known how each of these levels of investigation are affected by the various methods of filtration, and consequently how the filtration method used may affect the conclusions that are drawn from these metrics.

In this chapter we investigate the effect of different filtering methods on commonly used network metrics, using the association network of a troop of wild olive baboons as a case study. Global network metrics, individual network positions, and the extent of network substructuring were compared across the unfiltered network and networks that were filtered: (i) until the network structure was significantly different from random, (ii) by median association strength, (iii) by mean association strength, (iv) until the giant component was close to breaking up, and (v) by including only preferential associations.

Methods

Data Collection and Calculation of the Association Index

Data were collected on a troop of wild olive baboons (*Papio anubis*) in Gashaka-Gumti National Park (GGNP), north-eastern Nigeria, over a one-year period from March 2009-March 2010. GGNP is on the margin of the distribution of baboons, and is a somewhat unusual study site for baboons, as it includes large areas of rainforest and is the wettest of all the baboon study sites (Higham et al., 2009). The 'Kwano' study troop has been studied continuously since 2000 (Sommer and Ross 2011). For the current study, baboons were fully habituated to human observers and could be followed at a 2-6m distance. All baboons were individually recognised by the observers. The Kwano troop had a mean group size of 34 individuals (range 31-37) during the study period; however, only adults and subadults that were present for the entire study period were included in the networks (*i.e.* 16 individuals). The Kwano troop forms a 'closed' social system, in which members are clearly identifiable, and membership is highly stable. Nevertheless, fission-fusion dynamics have been observed in this troop (Alberts 2012); therefore, individuals do not always associate simultaneously with all troop members, but instead the troop temporarily splits into smaller subgroups.

Data were collected each day over an eight-hour period (*i.e.* 06:00 – 14:00 or between 10:00 – 18:00). Instantaneous sampling of the group, or scan sampling (Altmann 1974), was conducted every hour. During scan sampling, the identity of each baboon in sight was recorded at a preselected moment in time (*i.e.* every hour). Individuals that were seen together in a scan were considered to be associated. For five minutes before each scan, researchers walked around the area to locate baboons. The definition of an association used here is thus broader than an association based on individuals being in visual contact. The method used here may be a more appropriate estimation of associations at this site; the terrain at GGNP is very uneven and large parts are forested, and therefore using a purely visual definition of associations may underestimate the number of individuals in a subgroup. Each day around eight scans were collected, making a total of 467 scans.

Individuals were considered to be in association if they were both observed in a scan. Scan data were used to calculate the Twice Weight Index (TWI) in SOCPROG 2.4 (Whitehead 2009). The TWI was calculated for each dyad as follows:

$$TWI = \frac{X}{X + Y_a + Y_b}$$

Where X is the number of times *a* and *b* were seen together, Y_a the number of times *a* was seen but not *b*, and Y_b the number of time *b* was seen but not *a* (Cairns and Schwager 1987).

Filtering of Networks

The TWIs were used to create six networks: one unfiltered, and five filtered. In the unfiltered network all TWIs were used directly. The second network was filtered until network structure was significantly different from random. To this end, the observed network was filtered by increments of 0.01 and dichotomised until the mean clustering coefficient and the mean geodesic were significantly different from random. These two metrics were calculated for the observed network and compared to the distribution of the test statistics for 50 Erdös-Rényi random graphs (Erdös and Rényi 1959). The mean clustering coefficient and mean geodesic were chosen as test statistics because they provide a measure of the average cohesion of a network, and are complementary; the mean geodesic is a global network measure, in that it considers paths over the whole network, whereas the clustering coefficient focuses more on local structures (Croft et al., 2008). Network structure was found to be significantly different from a random structure at a filtration level of 0.170. This filter was applied to the weighted network; thus associations weaker than 0.170 (*i.e.* TWI < 0.170) were excluded. The third network was filtered by median association strength (0.220), and the fourth by mean association strength (0.224) (*i.e.* TWIs less than these values were excluded). The fifth network was filtered until its giant component was close to breaking up into smaller components. The giant component, or largest connected component, is the component in the network that contains the majority of nodes. The mean degree of a network is often close to 1 when the giant component is close to breaking up (Croft et al., 2008), and drops below 1 when the component fragments. Individuals dropping out of the network were not considered separate components, and thus only when a second component of at least two nodes was observed, was the network considered to have fragmented into multiple components. To determine the appropriate level of filtering using this method, the unfiltered network was visualised using NetDraw 2.089 (Borgatti 2002). The network was then filtered at increments of 0.0001, and was visualised again at each stage. These steps were repeated until the giant component was observed to break up (*i.e.* when more than one component was observed), and then the previous increment was used as a cut-off value. The cut-off value for this method was set at TWI < 0.340. Finally, the sixth network was filtered to include only preferential associations. Preferentially associating dyads are pairs of baboons that were observed to be in association significantly more frequently than expected by chance given the number of times those individuals were observed. We tested for preferential associations using SOCPROG 2.4 (Whitehead 2009). In this procedure, observed TWIs are compared to the distribution of TWIs of randomised data sets, in which data are randomised using a modification of the methods of Manly (1995) and Bejder, Fletcher, and Bräger (1998) (20,000 permutations), keeping group size and the number of times each individual was observed constant (Whitehead 1999). Dyads that were observed to associate significantly *more* than expected by chance were included in the weighted network.

Calculation of Global Metrics, Individuals Network Positions, and Substructures

All network metrics were calculated in UCINET (Borgatti et al., 2002). To test the effect of filtering on global network structure, we compared eight network metrics that are commonly used in the study of animal social networks: the binary and weighted density, binary mean degree, the largest connected component, the mean geodesic, the diameter, the mean clustering coefficient, and the weighted network centralisation (see Table 1 for definitions) across the six networks. These eight metrics are frequently used by researchers of animal networks to characterise the structure of relationships in a group. Networks were dichotomised to calculate the binary density and the binary mean degree. While all other metrics were calculated in weighted networks, only the weighted density and weighted network centralisation take the weights of the ties into account specifically. When a network had more than one component, the diameter was calculated as the largest geodesic observed in any component, whereas the clustering coefficient calculations only include individuals with more than one tie.

Table 1. Definitions of global network metrics (adapted from Wasserman and Faust 1994; Croft et al., 2008; Opsahl and Panzarasa 2009).

Global network metrics	
Binary density	The density of a network is a measure of the number of ties in a network, and indicates the level of cohesion. It indicates the number of ties in relation to the possible number of ties. For an undirected network: $$\Delta = \frac{E}{n(n-1)/2}$$ Where E is the number of ties in the network, and n the number of nodes. The value of Δ ranges between 0 (empty) to 1 (completely connected).
Weighted density	The weighted density of a network is a measure of the average weight of ties across all possible ties, and indicates how strongly connected a network is. $$\Delta_\omega = \frac{E_\omega}{n(n-1)/2}$$ Where $E\ \omega$ is the sum of the values of all ties, and n the number of nodes.
Binary mean degree	The mean degree is a measure of how well connected a network is. It indicates how many ties nodes in the network have on average. $$k = \frac{1}{n}\sum_i k_i$$ Where k_i is the number of nodes i is connected to. Higher values indicate that on average individuals have more ties.
Largest connected component	The largest connected component is the size of the largest group of nodes that are all reachable from each other.
Mean geodesic	The average shortest path length is a measure of how close, on average, two individuals are to each other in the network. It indicates the shortest path from a node to all other nodes in the network. $$L = \frac{1}{n}\sum_{i=1}^{n} d_{ij}$$ Where n is the number of nodes in the network, and d is the shortest distance between node i and j. Larger values indicate a greater distance between individuals and thus that relationships are less direct.
Diameter	The network diameter is the largest of the shortest paths between individuals in a network. This gives an indication of how 'wide' a network is, in other words, the maximum distance between nodes. Higher values indicate individuals in the network may be more distant from each other.

Global network metrics

Mean clustering coefficient	The mean clustering coefficient is a measure of the cliquishness of a network. It indicates the average proportion of ego's neighbours that are also connected to each other. $$C = \frac{1}{n}\sum_{i=1}^{n}\frac{2t_i}{k_i(k_i-1)}$$ Where t_i is the number of triangles of which node i is part, and k is the number of nodes i is connected to. The clustering coefficient ranges between 0-1, with large values indicating a large proportion of a node's neighbours also have ties between themselves (clustered).
Weighted network centralisation	The weighted network centralisation is a measure of how evenly ties are distributed over individuals in the network. It indicates the differences between the largest individual centrality score and the scores of all the other individuals in the network, and is normalised by maximum possible difference. $$C_D = \frac{\sum_{i=1}^{n}[C_D(n*) - C_D(n_i)]}{max\sum_{i=1}^{n}[C_D(n*) - C_D(n_i)]}$$ Where $C_D(n_i)$ is the centrality score for node i, and $C_D(n*)$ is the largest observed value. Low values indicate ties are equally distributed over individuals, high values indicate that a few individuals have most of the ties.

Next, we calculated four standard individual centrality measures frequently used in studies of animal social networks: degree centrality, betweeness centrality, closeness centrality and eigenvector centrality (see Table 2 for definitions). These centrality measures are frequently used by researchers of animal networks to identify animals that play important roles in their networks, and to draw conclusions about the characteristics of animals (*e.g.* age or sex) that may influence the role of individuals in their networks. Because centrality measures were calculated in weighted networks, the degree centrality and the eigenvector centrality use the sums of the values of an individual's ties. While betweeness and closeness centrality are distance-based measures, these do not take into account the weights of the ties, despite being calculated on a weighted network. To determine if individuals had similar network positions in networks that were filtered by different methods, and thus to assess the effect of filtering on the conclusions drawn about the roles of individuals in their networks, correlations were run between individual centrality scores of the same centrality measure across the networks. As data were not normally distributed, non-parametric correlations were used. Two sets of correlations were carried out; first, correlations were run including centrality scores of all individuals. However, researchers often draw conclusions on individual positions only within the connected component of the network, and correlations may be heavily affected by the inclusion of the centrality scores of the isolates. We therefore ran a second set of correlations including only those individuals that had a centrality score above zero. Finally, the presence and number of two types of commonly used substructures, cliques and *k*-plexes, were investigated in the weighted networks. The presence of substructures is frequently used by researchers of animal networks to identify layers in the social organisation of an animal group, and to identify individual characteristics that may underlie the formation of such layers (*e.g.* age-mates or kin may form clusters in a network). First, we determined the number of cliques, or maximally complete subgraphs, in the networks. We set the minimum size of cliques to three, as this is the smallest possible group above a dyad. Second, we searched for the number of *k*-plexes in the networks.