



OPEN

# A multilevel statistical toolkit to study animal social networks: the Animal Network Toolkit Software (ANTs) R package

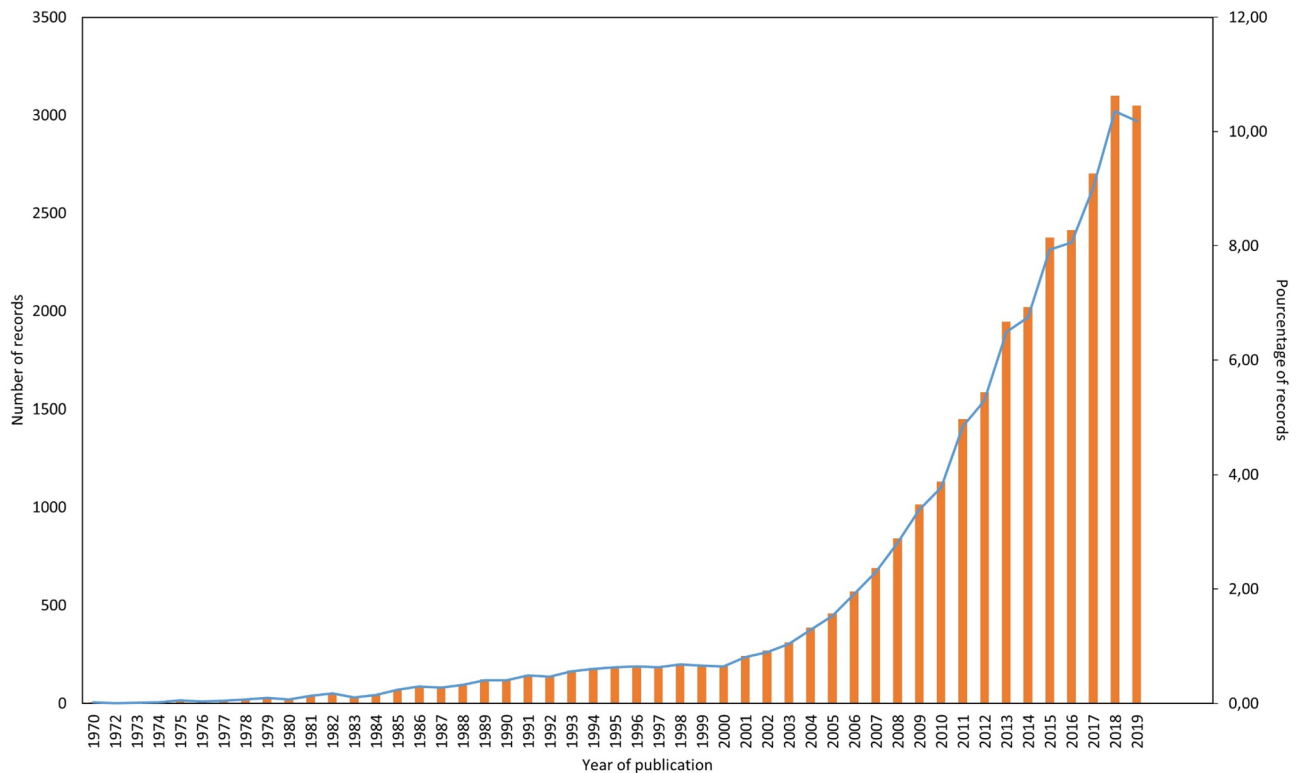
Sebastian Sosa<sup>1,2,6</sup>✉, Ivan Puga-Gonzalez<sup>3,6</sup>, Fenghe Hu<sup>4</sup>, Jérôme Pansanel<sup>2</sup>, Xiaohua Xie<sup>4</sup> & Cédric Sueur<sup>2,5</sup>

The possible role played by individual attributes, sociodemographic characteristics and/or ecological pressures in the interaction between animals and the development of social relationships between them is of great interest in animal ecology and evolutionary biology. Social Network Analysis is an ideal tool to study these types of questions. The Animal Network Toolkit Software (ANTs) R package was specifically developed to provide all the different social network analysis techniques currently used in the study of animal social networks. This global package enables users to (1) compute global, polyadic and nodal network measures; (2) perform data randomisation: data stream and network (node and link) permutations; (3) perform statistical permutation tests for static or temporal network analyses, and (4) visualise networks. ANTs allows researchers to perform multilevel network analyses ranging from individual network measures to interaction patterns and the analysis of the overall network structure, and carry out static or temporal network analyses without switching between different R packages, thus making a substantial contribution to advances in the study of animal behaviour. ANTs outperforms existing R packages for the computation speed of network measures and permutations.

In recent years the use of network theory has developed considerably in animal research (Fig. 1), facilitating the study of relationships between individuals and in a variety of research fields, including genetic<sup>1</sup>, spatial movements<sup>2</sup>, animal groups and species assemblages<sup>3</sup>, epidemiology<sup>4</sup>, information transmission<sup>5</sup>, ethology, sociobiology<sup>6</sup>, evolution<sup>7</sup>, conservation<sup>8</sup> and animal management<sup>9</sup>. Animal social networks are also dynamic as they change through time based on group composition and social choices/assemblages of individuals<sup>10–13</sup>. In addition, network properties may, as feedback, influence individuals' health or fitness through social transmission of disease and/or information (diffusion processes) leading to changes in social relationships and, consequently, to changes in the network topology<sup>14</sup>. SNA allows the study of all these aspects and has revolutionised our understanding of complex intra- and interspecific interactions between individuals. However, while methods have been specifically developed, none of the current software allows a unique environment to handle animal social networks analysis.

A social network is composed of nodes (e.g. individuals, groups, populations or species) and edges that represent spatial associations or any type of social relationships (i.e. the total number of affiliative or agonistic interactions over a given period of time). Regarding network analyses of group living animals, specific SNA have been developed to manage intrinsic biases according to the type of protocol used for data collection (focal sampling, scan sampling, group follow), the heterogeneity of individuals' observation and the group composition<sup>15,16</sup>. Given the inherent non-independence of data, the development of different network permutation approaches has permitted the creation of Null Models against which statistical hypotheses can be tested<sup>15</sup>. Although this progress has resulted in a tremendous diversity of SNA software, very few of these software propose an all-in-one

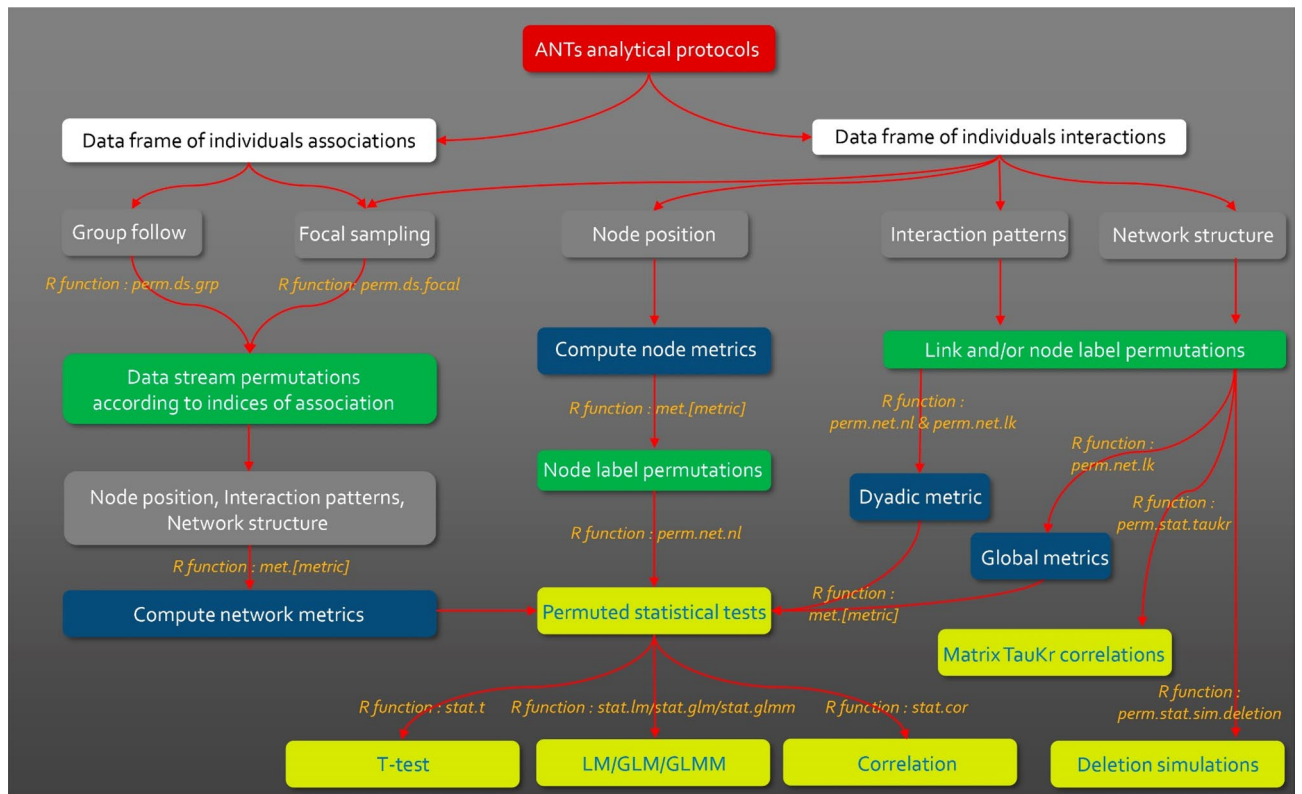
<sup>1</sup>Primate and Evolution Anthropology Laboratory, Anthropology Department, Sun Yat-Sen University, Guangzhou, China. <sup>2</sup>Université de Strasbourg, CNRS, IPHC UMR 7178, 67000 Strasbourg, France. <sup>3</sup>Institute for Religion, Philosophy and History, University of Agder, Kristiansand, Norway. <sup>4</sup>School of Data and Computer Science, Sun Yat-Sen University, Guangzhou, China. <sup>5</sup>Institut Universitaire de France, Paris, France. <sup>6</sup>These authors contributed equally: Sebastian Sosa and Ivan Puga-Gonzalez. ✉email: s.sosa@live.fr



**Figure 1.** Record count and percentage of records of the keywords “social networks animals” per year. Source: Web of knowledge.

toolbox for animal research. Even the most complete software programs, such as UCINET<sup>17</sup> and SOCPROG<sup>18</sup>, do not provide the most commonly used analysis procedures. Whilst the former lacks data stream permutations, the latter cannot be used for node label permutations and neither software offers time-aggregated SNA or has the high data flexibility of the R environment. R packages, on the other hand, offer high data flexibility, but there is no one package offering an all-in-one toolbox for animal research. For instance, ‘igraph’<sup>19</sup> and ‘sna’<sup>20</sup> are among the most used R packages. These are generalist packages (used for all types of network analyses) and therefore do not manage procedures used in animal social network research such as permuted statistical tests. Other packages, such as *asnipe*<sup>21</sup>, do integrate some data stream permutations but do not allow users to compute SNA measures or run statistical tests. Finally, *netTS* package allows users to run analyses on time-aggregated networks but does not include the standard procedure for analysing animal networks (node label and data stream permutations). The use of different R packages requires data to be manipulated and/or formatted in specific ways, which requires a certain level of programming knowledge. This switching between the different software packages and the absence of guidelines may cause users to make errors. Hence, these limitations may preclude the use of existing animal social network tools.

The creation of the Animal Network Toolkit software (ANTs) is an attempt to manage these limitations. Whilst ANTs may be of use for all scientists working on social network analyses, it was developed specifically for animal behaviourists and ecologists who want to apply SNA to their research questions with analytical protocols that have been thought, designed and applied for more than 20 years in these specific research areas. This was the main reason behind the publication of this paper and the instructions to use the package (analyses and indices implemented in the packages). This R package has been developed for researchers who are studying animal social networks (usually small networks of up to 1,000 nodes, although the package can be used for larger networks) and have specific observational data protocols. It provides them with a unique software to facilitate the processing of raw data and help researchers to select the most appropriate network measure and permutation approach for their data type, research question and run permuted statistical analyses to perform static and/or temporal (time-aggregated) network analyses. All these steps have been tailored in generic functions based on standard analytical procedures to offer an all-in-one toolbox that is similar to UCINET and SOCPROG, but has the flexibility of the R environment, thus allowing biologist non-expert in graph theory to follow specific analytical protocols, whilst expert users can use ANTs functions to speed up their specific analytical protocols (Fig. 2). Furthermore, ANTs computing speed outperforms ‘igraph’ for the SNA measures that are common to both packages and ‘asnipe’ for data stream permutations (see benchmarks in ESM Appendix 1, Table 1 to Table 6). The efficient and rapid performance of these processes is crucial because some data stream permutation analyses require SNA measures to be computed thousands of times. Finally, ANTs provides multiple variants of a single network measure, as well as R documentation with detailed explanations on these variants and their interpretations. While some of these variants appear in ‘igraph’ and ‘sna’ R packages, others are missing (e.g. Newman 2001 algorithm for shortest paths and related measures: diameter, betweenness, global efficiency). ANTs aims



**Figure 2.** ANTs analytical protocols. In white the types of data handled by ANTs, in gray the theoretical approaches and the types of data collections handled by ANTs, in green the family functions of class ‘perm.’ to perform data permutations, in blue the family functions of class ‘met.’ to compute network measures (node, polyadic or global), in yellow the family functions of class ‘stat’ to perform permutated statistical tests.

to provide users with an overview of each existing variant measure, thus enabling them to decide which one is the most appropriate for their research question<sup>22</sup>.

This paper outlines the large spectrum of analytic opportunities provided by ANTs:

- (1) Assess data collection reliability.
- (2) The calculation of the most commonly used and well-known network measures in animal studies as well as other, less well-known measures (Table 1).
- (3) A multilevel approach through the study of node, polyadic and global network measures.
- (4) Appropriate randomisation techniques (pre- or network permutations).
- (5) Several statistical tests (correlation test, t-test, General Linear Model, General Linear Mixed Model) with the calculation of permuted p-values, mean and CI posterior distribution.
- (6) The possibility to analyse static or temporal networks/comparative analyses for the comparison of different datasets (groups or species).
- (7) Network visualisation.
- (8) A fast computing process.

ANTs is currently available in an open beta test version on the following GitHub web page:

<https://github.com/SebastianSosa/ANTs> (for the stable release)

<https://github.com/ants-project/ANTs> (for the version under development).

## Methods

**Data input.** ANTs can process two types of data: (1) data representing the directed interactions of individuals (e.g. grooming) or their associations (e.g. proximity), and (2) data representing individual attributes (sex, age, dominance rank, etc.).

Interactions and association data can be input in the form of a matrix or a data frame(s). The data frame structure depends on the type of protocol the user wants to follow. For network permutations, data frames must be in an edge list format with at least two columns, one of which indicates the actor and the other the receiver. An additional column may indicate weights of interactions. These data frames allow the user to directly input data collected in the field. For the data stream permutation approach, data can be presented in data frame format. In this case, data frames are not edge lists because they contain additional information in extra columns. For data stream permutations concerning focal observations<sup>16</sup>, i.e. data obtained from following a specific individual over a certain amount of time<sup>23</sup>, two extra columns are required in addition to those indicating the givers and

Measure levels	Measures	Definition
Node	Degree	Number of links
	Outdegree	Number of outgoing links
	Indegree	Number of incoming links
	Strength	Sum of the weights of links
	Outstrength	Sum of the weights of outgoing links
	Instrength	Sum of the weights of incoming links
	R-index	Ratio of outgoing and incoming links
	Disparity	Variation of the weights of the links of a node
	Eigenvector	Node centrality considering the connections of a node and its alters
	Reach	Node centrality considering the overall strength of a node alters
	Affinity	Node centrality considering the overall strength of a node and that of its alters
	Laplacian centrality	Importance of a node in network cohesion
	Betweenness	Number of times a node lies within the shortest paths <sup>a</sup> of all pairs of nodes in the network
Polyadic	Assortativity <sup>b</sup>	Degree to which a node is linked to others with the same attribute rather than linked to others with a different attribute
	Geodesic distance	Shortest paths <sup>a</sup> between each pair of nodes in the network
Global	Density	The number of actual connections divided by all potential connections in the network
	Diameter	Greatest shortest path <sup>a</sup>
	Global efficiency	Efficiency of information transmission within the network
	Centralization index	Sum of the differences between the value of the maximum eigenvector centrality with a star network and the eigenvector centrality of each node in the real network <sup>c</sup>

**Table 1.** ANTs measures available and their short descriptions. <sup>a</sup>Shortest distance (based on the number of links or weights of the links) connecting a given pair of nodes. <sup>b</sup>The calculation of this measure is based on dyadic links but the value returned applies to the whole network.

receivers of a behaviour: one indicating the focal individual and another indicating the corresponding focal session. For data stream permutations on group follow observations<sup>5</sup>, i.e. recording individual associations at specific locations and time, data frames have to be in a ‘linear mode’, identical to SOCPROG (i.e. in which each line corresponds to the observation of an individual) with additional columns indicating the different ‘control’ factors (see “Permutations” section) such as the date, time of day or the geographical location associated with the interaction occurred.

It is also possible to use data frames for individual attributes (sex, age, dominance rank, hormone levels, etc.). These must be in a data frame format, with a row for every individual present in the data of individual interactions or associations. Each line represents the attribute(s) of a single individual.

Inputting these two types of data (interactions/associations and individual attributes) may enable the user to (1) permute and/or compute network measures on data representing individuals’ interactions or associations and (2) store node network measures with ANTs functions in the data frame(s) of individual attributes. This makes it possible to study how these node network measures are related to individuals’ attributes.

When performing the multiple networks analytical protocol, the user has to create an R list object where each element of the list stores interaction/association data representing a single network (list of data of interactions or associations). This list must contain a unique data format of interactions/associations (i.e. only edge lists, associations of group follow or associations of focal sampling). Optionally, the user can create a second R list object with the attributes of the individuals present in the corresponding list of interactions/associations (e.g. the data frame of individual attributes in element 1 corresponds to the individuals present in the list of interactions/associations in element 1, etc.). This way, permutations are generated independently in each network (e.g. 1,000 permutations in network 1, 1,000 permutations in network 2, etc.).

**Testing data collection robustness.** One of the main issues with regard to social network analysis and the study of animal groups is the quality of data collection (time of observation), as observation biases (e.g. some individuals are more frequently observed than others) can generate unreliable statistical results<sup>24, 25</sup>. Usually, data collection protocol has to be planned for the needs of the intended SNA before collecting data. The following questions must be answered: Do I observe all group members equally? Am I using the best method to limit the disturbance of animal behaviour and interactions? The choice of observation period is also a key factor, as some interindividual associations or interactions are rare and/or difficult to observe over the short term but are still important to attain the objectives of the study. However, this not always the case as scientists often collect data before carrying out analyses. ANTs meets the needs of these differing approaches by offering two different protocols to assess data collection robustness:

1. Lusseau, et al.<sup>24</sup> protocol to assess the robustness of node measures through bootstrapping.
2. Balasubramaniam, et al.<sup>25</sup> protocol to assess the robustness of global measures through observation deletion simulations.

For further information on the use of these different protocols, please refer to ANTs R documentation concerning functions in the ‘sampling.’ family.

**Controlling for time heterogeneity.** It is sometimes difficult to obtain the same number of observations per individual. ANTs enables users to control for time heterogeneity in different ways through the use of different association indices, namely the generalised affiliation index, the simple ratio index, the half-weight index or the square root index<sup>6</sup>. For further instructions on the use of these different indices, please refer to ANTs R documentation concerning the functions in the ‘assoc.’ family.

**Computing network measures.** Three types of network measures can be identified depending on the level of organisation: global measures, polyadic measures, and node measures. In ANTs, all these measures are grouped under the function family ‘*net*’. All the node measures available in ANTs are synthesised in Table 1. The measures we proposed in the package ANTs are the ones commonly used in Animal Social Network Analyses<sup>6, 22, 26–28</sup>.

Global measures (e.g. network diameter) are used to study the overall network and obtain valuable information regarding network efficiency, resilience, clusterisation, etc. Polyadic measures (e.g. assortativity) allow the study of interaction patterns between individuals. These measures provide information about how individuals interact according to their attributes. Node measures (e.g. strength) are the most frequently used measures in animal research. Among other things, node measures inform users about the centrality of an individual, the number of alters it has and/or its activity according to individual attributes, and reveal patterns that are common to individuals with similar attributes. By giving access to global, polyadic and node measures, we aim to enable users to adopt a multilevel approach and thereby understand the centrality of individuals in a group, the patterns of interaction between them and the impact of these two levels on the global network structure<sup>22, 29</sup>.

For more details on the different types of measures, their mathematical formula, interpretation, limitations and past use in animal research, see Whitehead<sup>6</sup>, Sueur, et al.<sup>26</sup>, Sosa, et al.<sup>22</sup>, Sosa<sup>29</sup> and refer to ANTs R documentation.

**Permutations.** When considering data robustness, permutations can be used to avoid observation biases and ensure the reliability of results obtained by SNA (i.e. results that have no type I and type II errors). Indeed, with the exception of some specific cases such as experiments in social insects, where individuals may be tracked continuously, it is usually assumed when examining inter-individual interactions within a group or a population that neither all the interactions nor all individuals are observed, that the times of observation vary from one individual to another, and that the data collected are intrinsically dependent. For these reasons, permutation tests are needed to control for data independency before performing inferential statistical tests, as inferential statistical tests assume data independency<sup>16</sup>.

The Null Model (NM) approach via permutation is one of the many current possibilities to test statistical hypotheses<sup>15</sup>. It allows users to perform analyses by creating random data sets from the observed data. The observed measure of interest  $X$  (e.g. coefficient of correlation) is compared to a posterior distribution obtained from the random data sets, and assesses whether  $X$  is significantly different from the random distribution by calculating the proportion of random values that differ from the observed value. The NM approach can be applied in different ways. ANTs allows for this by adapting the permutations (pre- or network permutations) according to the type of data collected (i.e. pre- or network permutations for data on associations and interactions respectively) and the research question (i.e. permuting nodes when examining individual network measures or permuting links when examining individual polyadic or global measures).

Data stream and node network permutations are two of the most commonly used permutation methods to build null models in animal social network analysis. A description of these methods is presented by Puga-Gonzalez et al. (submitted). Data stream permutations were initially used to test whether individuals in a social population have a preference for association with certain partners rather than with others<sup>27, 30</sup>. One of the advantages of this method is that it can control for different factors such as location. It is therefore possible to test whether non-random associations are due to individuals’ social preference or result from a preference for the same habitat or location<sup>27</sup>.

Node network permutation is the other commonly used method to test network-related hypotheses in animal research. Node permutations have mainly been used to compare two matrices (or networks) involving the same group of individuals, i.e. matrix correlations. In this case, the values entered in the cell of the matrices are (un) directed behaviours (e.g. grooming or playing). In contrast to the gambit of the group, (un)directed behaviours are usually collected via focal sampling, scan sampling, or ad libitum sampling<sup>23</sup>. During node permutations, the identity of the nodes is redistributed at each permutation whilst the node metric is kept constant. This allows users to test whether a specific network metric is associated with a specific node attribute (e.g. whether females groom more than males), or whether behaviours are reciprocated or directed to individuals with a specific trait (e.g. grooming directed up the dominance hierarchy). All of the permutation approaches available in ANTs are in the family function ‘*perm*’ with two subclasses, ‘*perm.ds*’ and ‘*perm.net*’ for data stream and network permutations, respectively. ANTs can perform data stream permutations for group follow and focal sampling data collection protocols. Network permutations can be performed on (1) node label(s) (with labels’ dependency maintained or not), (2) links, (3) link weights, and (4) link weights swap between categories. Among those different types of permutations, node label (ESM Appendix 1) and data stream (ESM Appendix 2) permutations are probably the most commonly used standard approaches in animal network analysis. For this reason, we developed a specific workflow to allow their use (ESM Appendix 1 and ESM Appendix 2) in ANTs for the study of single<sup>31</sup> or multiple networks<sup>9, 13</sup> (for network comparisons or time-aggregated analyses). To date, ANTs is

the only software permitting the use of these approaches in an all-in-one environment and their application for the analysis of multiple networks.

For more details on the different permutations and their applications according to the data collection protocol, the type of behavioural data collected and the research question, see Bejder, et al.<sup>30</sup>, Whitehead<sup>27</sup>, Whitehead, et al.<sup>32</sup>, Croft, et al.<sup>28</sup>, Farine<sup>16</sup>, Momigliano, et al.<sup>33</sup>, Sosa<sup>29</sup>, ANTs R documentation, ESM Appendix 1 and ESM Appendix 2.

**Statistical tests based on data permutations.** All the statistical tests available in ANTs are in the family function ‘stat’. The available tests are correlation test ‘*stat.cor*’, t-test ‘*stat.t*’, Linear Model (LM) ‘*stat.lm*’, Generalised Linear Model (GLM) ‘*stat.glm*’, Generalised Linear Mixed Models (GLMMs) ‘*stat.glmm*’, assortativity test ‘*stat.assortativity*’, TaurK correlation ‘*stat.Taurk*’ and deletion simulation ‘*stat.deletion*’. ANTs *stat.* function returns an object with the posterior distribution of the variable tested.

1. Once the permutation test has been performed, the function ‘ant’, allows the user to obtain the statistical results from any output object of any function ‘stat’. The ‘ant’ function returns a data frame with statistics specific to the type of statistical test run. However, some of these statistics are common to all tests, namely the P-values on the right or left of the distribution and the two-side p-values.
2. Measures of the ‘effect size’ of the posterior distribution according to the statistics of interest: 95% confidence interval and the mean of the distribution a posteriori (see Farine and Whitehead<sup>34</sup>). The histograms of the post-distribution of the statistics of interest obtained from the permutations.

**Network visualization.** ANTs allows network visualisation with a data frame containing node information and a matrix of interactions/associations. Nodes and links can be parametrised to modify their size and colour and highlight differences (e.g. females showing higher eigenvectors than males). Network layouts are currently based on Barnes Hut repulsion, Hierarchical Repulsion and Force Atlas 2. For more details on network visualisation, see ANTs function ‘*net.vis*’ in the package instructions document. These layouts are commonly used in animal social network analyses<sup>9, 35–37</sup> as for instance, Force Atlas 2 arranges the visualisation graph with the distance between nodes is inversely proportional to their association, giving a nice view of who is close to whom.

## Discussion

ANTs provides researchers with an all-in-one software containing the most commonly used analytical tools in research on animal social networks. In order to offer tailored functions, ANTs provides what we call analytical protocols that represent standard procedures developed in animal social network analysis over the past decade, namely (1) node label permutations, (2) data stream permutations (for group follow and focal sampling). These two analytical protocols can be used on single networks or time-aggregated networks, allowing unexperienced users to run complex analyses on social networks. Besides, ANTs is also a flexible tool that enables users to conduct each step independently, thus providing users who have programming and analytical skills with an ideal tool to run specific analyses according to their requirements. This will enable users to focus less on coding and more on the selection of the most appropriate permutation approach, statistical test(s) and measures(s), and indeed consider in greater detail why one measure (or version thereof) should be chosen rather than another<sup>22</sup>. ANTs functions ‘*which.metric*’ and ‘*which.protocol*’ might be used for an interactive selection. The current version of ANTs includes the most common SNA tools for animal research. New network measures and analytical protocols will be added in the near future with the same objectives: (1) introduce all the variants of each measure, as they are likely to have an impact on the biological interpretations, (2) explain the interest of the measures and protocols depending on the biological question, and (3) optimise the computation speed of the functions. We consider this package to be upgradeable and collaborative and planning to see new indices implemented in ANTs such as those described in these two perspectives papers<sup>18, 38</sup>.

Received: 21 November 2019; Accepted: 2 July 2020

Published online: 27 July 2020

## References

1. Brent, L. J. *et al.* Genetic origins of social networks in rhesus macaques. *Sci. Rep.* **3**, 1042 (2013).
2. Mourier, J., Ledee, E. J. & Jacoby, D. M. A multilayer perspective for inferring spatial and social functioning in animal movement networks. *BioRxiv*:749085 (2019).
3. Pasquaretta, C. *et al.* A spatial network analysis of resource partitioning between bumblebees foraging on artificial flowers in a flight cage. *Mov. Ecol.* **7**, 4 (2019).
4. Duboscq, J., Romano, V., MacIntosh, A. & Sueur, C. Social information transmission in animals: Lessons from studies of diffusion. *Front. Psychol.* **7**, 1147 (2016).
5. Allen, J., Weinrich, M., Hoppitt, W. & Rendell, L. Network-based diffusion analysis reveals cultural transmission of lobtail feeding in humpback whales. *Science* **340**, 485–488 (2013).
6. Whitehead, H. *Analyzing Animal Societies: Quantitative Methods for Vertebrate Social Analysis*. (University of Chicago Press, 2008).
7. Sueur, C., Romano, V., Sosa, S. & Puga-Gonzalez, I. Mechanisms of network evolution: A focus on socioecological factors, intermediary mechanisms, and selection pressures. *Primates* **60**, 167–181 (2019).
8. VanderWaal, K. L., Atwill, E. R., Isbell, L. A. & McCowan, B. Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biol. Cons.* **169**, 136–146 (2014).
9. 9SOSA, S. O. *et al.* Impact of group management and transfer on individual sociality in Highland cattle (*Bos taurus*). *Front. Vet. Sci.* **6**, 183 (2019).

10. 10Boucherie, P. H., Sosa, S., Pasquaretta, C. & Dufour, V. A longitudinal network analysis of social dynamics in rooks *Corvus frugilegus*: Repeated group modifications do not affect social network in captive rooks. *Curr. Zool.* zow083 (2016).
11. 11Borgeaud, C., Sosa, S., Bshary, R., Sueur, C. & Van de Waal, E. Intergroup variation of social relationships in wild vervet monkeys: a dynamic network approach. *Front. Psychol.* 7 (2016).
12. 12Sosa, S., Zhang, P. & Cabanes, G. Social networks dynamics revealed by temporal analysis: An example in a non-human primate (*Macaca sylvanus*) in “La Forêt des Singes”. *Am. J. Primatol.* 79 (2017).
13. Borgeaud, C., Sosa, S., Sueur, C. & Bshary, R. The influence of demographic variation on social network stability in wild vervet monkeys. *Anim. Behav.* 134, 155–165 (2017).
14. Croft, D. P. *et al.* Effect of gyrodactylid ectoparasites on host behaviour and social network structure in guppies *Poecilia reticulata*. *Behav. Ecol. Sociobiol.* 65, 2219–2227 (2011).
15. Croft, D. P., Madden, J. R., Franks, D. W. & James, R. Hypothesis testing in animal social networks. *Trends Ecol. Evol.* 26, 502–507 (2011).
16. 16Farine, D. R. A guide to null models for animal social network analysis. *Methods Ecol. Evol.* (2017).
17. 17Borgatti, S. P., Everett, M. G. & Freeman, L. C. *Ucinet for Windows: Software for Social Network Analysis.* (2002).
18. Díaz, J., Petit, J. & Serna, M. A survey of graph layout problems. *ACM Comput. Surv. (CSUR)* 34, 313–356 (2002).
19. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *Int. J. Complex Syst.* 1695, 38 (2006).
20. Stoffel, M. A., Nakagawa, S. & Schielzeth, H. rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods Ecol. Evol.* 8, 1639–1644 (2017).
21. Fowler, J. H., Dawes, C. T. & Christakis, N. A. Model of genetic variation in human social networks. *Proc. Natl. Acad. Sci.* 106, 1720–1724 (2009).
22. 22Sosa, S., Sueur, C. & Puga-Gonzalez, I. Network measures in animal social network analysis: Their strengths, limits, interpretations and uses. *Methods Ecol. Evol.* <https://doi.org/10.1111/2041-210x.13366>.
23. Altmann, J. Observational study of behavior: Sampling methods. *Behaviour* 49, 227–266 (1974).
24. 24Lusseau, D., Whitehead, H. & Gero, S. Incorporating uncertainty into the study of animal social networks. arXiv preprint arXiv:0903.1519 (2009).
25. 25Balasubramaniam, K. N. *et al.* The influence of phylogeny, social style, and sociodemographic factors on macaque social network structure. *Am. J. Primatol.* (2017).
26. Sueur, C., Jacobs, A., Amblard, F., Petit, O. & King, A. J. How can social network analysis improve the study of primate behavior?. *Am. J. Primatol.* 73, 703–719 (2011).
27. Whitehead, H. Testing association patterns of social animals. *Anim. Behav.* 57, F26–F29 (1999).
28. Croft, D., James, R. & Krause, J. *Exploring Animal Social Networks* (Princeton University Press, Princeton, 2008).
29. 29Sosa, S. in *Encyclopedia of Animal Cognition and Behavior* (eds Vonk, J. & Shackelford, T.) 1–18 (Springer, New York, 2018).
30. Bejder, L., Fletcher, D. & Bräger, S. A method for testing association patterns of social animals. *Anim. Behav.* 56, 719–725 (1998).
31. 31Sosa, S. The influence of gender, age, matriline and hierarchical rank on individual social position, role and interactional patterns in *Macaca sylvanus* at ‘La Forêt des singes’: A multilevel social network approach. *Front. Psychol.* 7 (2016).
32. Whitehead, H., Bejder, L. & Ottensmeyer, C. A. Testing association patterns: Issues arising and extensions. *Anim. Behav.* 69, e1 (2005).
33. 34Momigliano, P. *et al.* Extraordinarily rapid speciation in a marine fish. *Proc. Natl. Acad. Sci.* 201615109 (2017).
34. Farine, D. R. & Whitehead, H. Constructing, conducting and interpreting animal social network analysis. *J. Anim. Ecol.* 84, 1144–1163 (2015).
35. 36Chapman, C. A. *et al.* Social behaviours and networks of vervet monkeys are influenced by gastrointestinal parasites. *PLoS one* 11 (2016).
36. Perryman, R. J. *et al.* Social preferences and network structure in a population of reef manta rays. *Behav. Ecol. Sociobiol.* 73, 114 (2019).
37. Schilds, A. *et al.* Evidence for non-random co-occurrences in a white shark aggregation. *Behav. Ecol. Sociobiol.* 73, 138 (2019).
38. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069 (2010).

## Author contributions

S.S. and P.-G.I. conceived the ideas, coded the main part of the software, and led the writing of the manuscript. F.H. coded part of the software on geodesic distance and diameter. J.P. coded part of the software on visualization and contributed critically to the drafts. X.X. contributed partially to the implementation of algorithms and critically to the drafts. S.C. conceived the ideas, contributed critically to the drafts. All authors gave their final approval for publication.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-69265-8>.

**Correspondence** and requests for materials should be addressed to S.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020