



HAL
open science

Conceptual representations of animal social networks: an overview

Matthew Silk

► **To cite this version:**

Matthew Silk. Conceptual representations of animal social networks: an overview. *Animal Behaviour*, 2023, 201, pp.157-166. 10.1016/j.anbehav.2023.04.017. hal-04154538

HAL Id: hal-04154538

<https://hal.science/hal-04154538>

Submitted on 6 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Conceptual representations of animal social networks: an overview

Matthew J. Silk^{1*}

1. CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

*corresponding author:

Email - matthewsilk@outlook.com

Phone - +44(0)7858292920 / +33(0)749875091

Address - 1919 route de Mende, 34090, Montpellier, France.

Word count: 9620

Abstract

Networks are now widely used to represent, quantify and model animal behaviour. These approaches have proved valuable in linking individual behaviours to emergent population-level patterns, and quantifying the implications of these population structures for wider ecological and evolutionary processes. However, there are diverse approaches available to represent network data and choosing the right tool to answer a particular question can be challenging. Here I provide an overview of different network representations, highlighting their potential applications in behavioural ecology and drawing attention to key resources to help with their implementation. My aim is to provide an accessible guide that helps behavioural ecologists take full advantage of the potential of the different ways in which their data can be used to generate social (and spatial) networks.

Keywords: social network; dynamic network; multigraph; hypergraph; simplicial complex; multilayer network; associations; social groups; social relationships

Highlights

- A guide of different approaches to represent animal social network data
- An overview of how to employ and when best to use different network representations
- Introduction to less used methods such as multigraphs, and signed and ego networks
- Case study showing association data as bipartite, dyadic and higher-order networks

Introduction

Social network analysis is now pervasive in behavioural ecology. It has helped us understand how behavioural interactions between individuals scale up to group- or population-level social structures. This has played a key role in helping answer key questions related to social ecology and evolution, for example being applied to study dominance and group stability (Shizuka & McDonald 2012; Dey & Quinn 2014; Silk *et al.* 2019; Hobson *et al.* 2021a), cultural evolution (Allen *et al.* 2013; Aplin *et al.* 2015; Kulahci *et al.* 2018), infectious disease transmission (Hamede *et al.* 2009; VanderWaal *et al.* 2014; Silk *et al.* 2018c; Powell *et al.* 2020), cooperation (Edelman & McDonald 2014; Dakin & Ryder 2020; Connor *et al.* 2022) and population responses to selection (Farine & Sheldon 2015; Fisher & McAdam 2017; Wice & Saltz 2021).

However, since its introduction as a tool the use of social network approaches have changed considerably over time alongside the questions they have been applied to (Pinter-Wollman *et al.* 2014; Croft *et al.* 2016; Cantor *et al.* 2021). Increasingly researchers have moved to consider dynamic networks that consider changes in social relationships over time (Farine 2017b; Fisher *et al.* 2017a) or multilayer networks that consider different types of interaction or association within the same network object (Silk *et al.* 2018b; Finn *et al.* 2019). More recently there has been a real push in the network science community to embrace network approaches that move beyond solely considering dyadic representations of interactions (Musciotto *et al.* 2022; Silk *et al.* 2022).

As a result, there are now diverse options available to an animal behaviour researcher when choosing an approach to represent their social system, even prior to selecting appropriate statistical analyses or modelling approaches. Here, instead of a detailed 'How To', I provide an overview of the principal ways to represent relational datasets in animal societies, pointing towards more detailed resources for less-used approaches and offering guidance on when different approaches are most valuable. Even for the same dataset it is possible to represent interaction or association data in different ways, meaning decisions about the best approach to use should depend first and foremost on the research questions of interest (Carter *et al.* 2015) and second on any limitations inherent to the dataset analysed. I hope to provide a resource that can be valuable for researchers designing animal social network projects and analyses (e.g. for pre-registered studies or funding applications) as well as for those faced with previously collected data with particular limitations or constraints.

The building blocks of networks for animal behaviour

The basics of how to construct networks in ecology (Proulx *et al.* 2005) and animal social behaviour (e.g. (Croft *et al.* 2008; Farine & Whitehead 2015)) have been discussed extensively elsewhere so I provide only a very brief overview that indicates the scope and focus of this paper.

Networks consist of nodes or vertices connected by links or edges between them. In most animal social networks nodes will represent individuals. However, nodes can also represent particular groups, locations (e.g. burrows, watering holes, food resources) or time-points that individuals are connected to. While not strictly social networks, movement networks in which nodes represent locations and edges the movements of individuals (Jacoby & Freeman 2016), also provide valuable tools to study animal social behaviour. Edges represent some form of interaction or association between the nodes. Examples in animal social behaviour include specific behavioural interactions (e.g. grooming, aggressive interactions), close contact (e.g. as recorded using proximity loggers) or co-occurrence of individuals within a group or at a particular location. Sometimes edges will represent more abstract indications of the social relationship between individuals based on one or more of these data sources.

In typical network approaches all nodes will represent the same “type” of entity (e.g. you wouldn’t mix individuals and locations) and edges only connect pairs of nodes (or dyads). However, alternative network representations also discussed below relax these restrictions. For example, bipartite networks can be used to represent networks with interactions *only between* two types of node (Larremore *et al.* 2014), multilayer networks allow the representation of different types of nodes and/or edges together, and higher-order network approaches allow for edges that connect more than two individuals together. As a result, the full network toolkit available to researchers provides considerable flexibility to researchers that can be adapted to the questions they have.

An important note here, is that regardless of the network representation used it is important that animal social network data is typically only a sample of the individual (nodes), associations/interactions that occur (edges) or both and it is important to account for this appropriately with subsequent analyses. This will be especially important when sampling is biased or uneven. How best to do this will vary considerably depending on the research questions being asked, the method of data collection and the representation being used and is beyond the scope of this paper. However, readers are encouraged to explore the recent literature discussing statistical methodologies to deal with these sampling issues in a range of contexts (Young *et al.* 2020; Franks *et al.* 2021; Hart *et al.* 2021; Ross *et al.* 2022).

Static networks

Binary networks

A binary, undirected network is the most basic representation of social relationships. Edges (or links) connect vertices (nodes) when two individuals share a meaningful social relationship. When measuring animal social networks these edges typically represent individuals detected in close proximity (e.g. contact networks: (White *et al.* 2017)), recorded in an aggregation or group together (Franks *et al.* 2010) or interacting in a specific way (e.g. grooming: (Cowl *et al.* 2020); fighting: (Hobson *et al.* 2021a)). A long history of researchers using binary, undirected networks means that there are a wealth of approaches available to study the position of individuals within the network, features intermediate between individual and group or population properties (mesoscale structure) such as cliques and communities, as well as properties of the network as a whole (see overviews provided by (Croft *et al.* 2008; Wey *et al.* 2008; Farine & Whitehead 2015; Silk *et al.* 2017b)). In addition, some measures calculated in more complex representations are inherently binary, such as degree (the number of connections an individual has). However, by representing social relationships as dyadic, unweighted and undirected (or by using binary measures) a considerable amount of information is typically lost meaning the use of unweighted networks in animal behaviour research is frequently cautioned against in most cases (see discussion and simulation studies in (Franks *et al.* 2010; Croft *et al.* 2011; Farine 2014; Farine & Whitehead 2015)), although can be useful occasionally as long as the threshold used is carefully justified (Croft *et al.* 2011). In R representation and basic analysis of binary networks is possible using the packages *igraph* (Csardi & Nepusz 2006), *sna* (Butts 2008a, 2014) and *network* (Butts 2008b, 2015) among others.

Weighted networks and multigraphs

Another easy way to add information to a network representation is to incorporate data on the frequency, duration or strength of social relationships. In weighted networks edges are assigned a value to indicate the strength of the connection between two individuals. Relatively early in the history of animal social network analysis, the importance of using weighted edges to represent interactions and associations was recognized (Franks *et al.* 2010; Farine 2014), and a wide variety of different ways of weighting edges are now available for researchers to use in different contexts, such as commonly-used association indices (Hoppitt & Farine 2018) or equivalents that retain information on the number of observations (Hart *et al.* 2021). Most common network analyses have now been generalized for weighted networks including many key network measures and community detection algorithms, typically using the same software packages highlighted in the previous section. In

addition, a number of statistical tools have recently been developed that deal with sampling issues in weighted network representations that offer potentially powerful tools in animal social network analysis (see approaches described in (Young *et al.* 2020; Hart *et al.* 2021; Ross *et al.* 2022)). However, it can be very important to consider how edge weights are included in these calculations and subsequent analyses. For example, when calculating the shortest paths through the network (and measures derived from this such as betweenness and closeness centrality) edges with higher weights could be considered “shorter” or “longer” depending on context. Moreover, how the path length calculated depends on edge weight need not always have the same mathematical form. A key example of this is that default behaviour of the R package igraph (Csardi & Nepusz 2006) treats edge weights as a cost (e.g. as if they are a distance) while in many applications in animal behaviour the opposite is true, meaning the edge weights used have to be redefined for many applications of these approaches in behavioural ecology.

What is often neglected in animal social network research, however, is that repeated interactions can also be represented as a multigraph, where there is no constraint on the number of edges connecting a dyad (Fig. 1b). In some contexts, this might be a more appropriate way of representing the frequency of interactions between pairs of individuals (e.g. A groomed B 10 times), and it certainly helps more intuitively generate reference models for subsequent analysis as edge rewiring algorithms (a permutation approach in which at each step the identity of nodes connected by a selected edge are changed) can be applied to each edge independently (Hobson *et al.* 2021b). In addition, for some animal network datasets including both the frequency and duration or strength of interactions could provide valuable additional information. In this case it is necessary to analyse a weighted multigraph (Fig. 1c). For example, when testing hypotheses in grooming networks it may be that the length of individual grooming bouts and frequency of grooming bouts provide independent information about the social relationship between two individuals. Taking this approach can open up new research questions about the relative importance of frequent versus long duration social interactions in group social structure and stability. It may also be important, for example, when studying social contagions if the likelihood of transmission depends on the duration of interactions (e.g. for less infective pathogens or for social learning of more complex behaviours). When using multigraphs in this way it will be important to control for variation in sampling effort (e.g. the number of times individuals were observed) if this varies among individuals when conducting subsequent statistical analyses. Using multigraphs can also create challenges when representing networks visually (especially with many edges between nodes) and is more computationally intensive than using weighted networks, so using them should be reserved for contexts in which maintaining the independence of the frequency and duration of interactions provides valuable

additional information on social relationships and their ecological and evolutionary consequences as outlined above. The R package igraph (Csardi & Nepusz 2006) facilitates storing and representing weighted multigraphs.

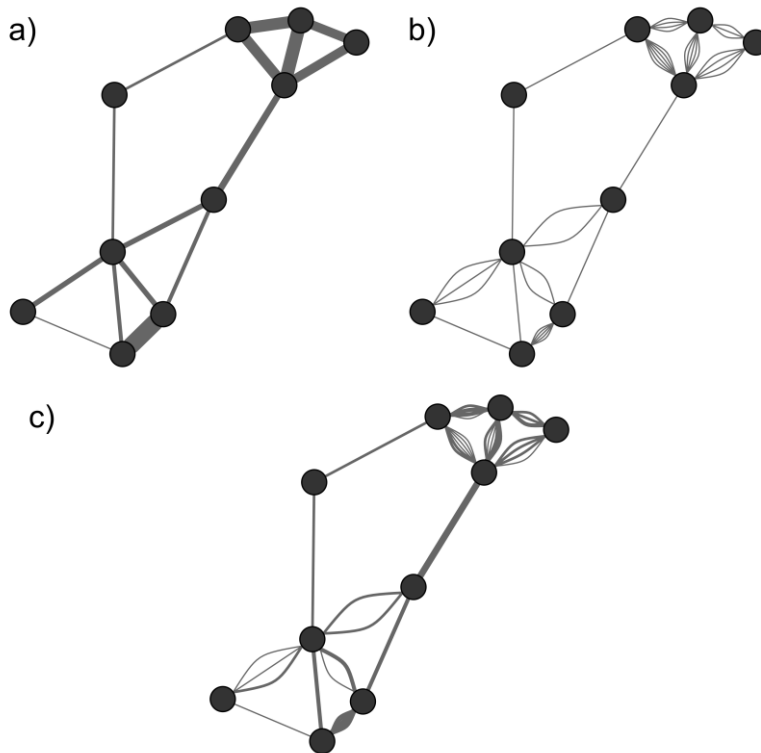


Figure 1. Examples of a) a weighted network in which edge width represents the total duration of interactions, b) an unweighted multigraph in which the number of edges in a dyad represents the number of interactions between two individuals, and c) a weighted multigraph in which the number of edges in a dyad represents the number of interactions between two individuals and the width of each edge illustrates the duration of each interaction. The weighted multigraph retains the most information about social interactions within the group.

Directed networks

Another simple way to include additional information is to use a directed network instead. In directed networks (or digraphs) edges no longer simply represent a connection between node A and node B, they indicate a connection *from* node A *to* node B. Directed networks have proven especially useful for representing behavioural interactions, in which who initiates an interaction is important. For example, directed networks have been commonly used to provide insights into the structure of dominance hierarchies (Dey & Quinn 2014; Silk *et al.* 2019; Hobson *et al.* 2021a). When studying hierarchies, considering directed edges makes it possible to distinguish between transitive (individual *i* is dominant over individual *j* and *k*,

individual j is dominant over individual k) and cyclical (individual i is dominant over individual k , individual j is dominant over individual k and individual k is dominant over individual i) triads (Shizuka & McDonald 2012; Dey & Quinn 2014) or investigate reciprocity effects (Dey & Quinn 2014; Silk *et al.* 2019) which are important indicators of the linearity and stability of dominance hierarchies. Directed edges have also proved valuable in studying grooming behaviour and affiliative interactions within groups (Balasubramaniam *et al.* 2018; Cowl *et al.* 2020). Finally, they have also provided valuable tools in studying movement behaviour, especially when it is possible to define discrete patches or locations that individuals move between (reviewed by (Jacoby & Freeman 2016)). Many common measures and tools have been extended to use in directed networks, and some additional measures are available (e.g. PageRank centrality; (Ding *et al.* 2009)). Directed networks can be represented and relevant measures calculated using the same software tools as binary and weighted networks. Similarly, Bayesian statistical models that are designed account for sampling issues common in empirical social network data can be applied to directed network data (see modelling frameworks provided by (Hart *et al.* 2021; Ross *et al.* 2022)).

Signed networks

While conventionally non-zero edges or links will have positive values (as for all the forms of network discussed so far), researchers in other fields have started to use signed networks (Beigi *et al.* 2016; Kirkley *et al.* 2019). In signed networks positive (e.g. affiliative) interactions are assigned positive edge weights and negative (e.g. agonistic, avoidance) interactions are assigned negative edge weights (although in practice most analysis of signed networks has focused on binary versions in which edge weights are either +1 or -1). Many applications of networks to study behavioural interactions in animal groups could lend themselves to the analysis of signed networks when there is information available about different types of interaction and these are considered to be either “positive” or “negative”. Perhaps the most direct application will be in comparing patterns of affiliative interactions (positive) with patterns of social avoidance (negative), especially given social avoidance is relatively understudied in non-human animals (Strickland *et al.* 2017). Another context in which signed networks may be valuable is in comparing networks of affiliative and agonistic interactions, although note that for some research questions it may be more effective to integrate these behaviours as a multilayer network (see below).

Signed network approaches are likely to be especially beneficial when studying questions related to group stability (e.g. (McCowan *et al.* 2011; Larson *et al.* 2018)) and alliance formation (e.g. (Connor *et al.* 2022)). Structural balance theory (Ilany *et al.* 2013), which predicts for example that if two individuals share a strong mutual connection they are unlikely to possess a negative social relationship, extends neatly into signed networks (see

Fig. 1 in (Facchetti *et al.* 2011)). The presence of unbalanced triads (e.g. the triad described in the previous sentence) in a signed network could be used to predict network dynamics or changes in group composition. There may be interest in investigating the relative importance of structural imbalances involving different combinations of positive and negative ties. Similarly, the commonly-used stochastic block model has been extended to signed networks with the aim of identifying clusters or communities characterized by positive interactions with each other but negative ties with other clusters (Jiang 2015). This offers a very natural way to characterize complex social structures such as those dominated by matrilineal or where alliances govern access to mates or reproductive opportunities. Analysis of signed networks in R can be conducted using the R package *signnet* (Schoch 2020), an extension of *igraph* for signed networks.

Bipartite (two-mode) networks

Bipartite (or two-mode) networks connect two different types of node, with connections between the same type of node not possible (Larremore *et al.* 2014) (Fig. 3). They have been commonly used to study ecological networks such as plant-pollinator (Dupont *et al.* 2014; Miele *et al.* 2020) or host-pathogen (Valverde *et al.* 2020; Albery *et al.* 2021) networks. However, one of the classic social network datasets – the Davis' Southern women network – illustrates the potential values of bipartite networks in behavioural ecology. The Southern women network links a set of 18 women to the parties or events that they attended (Davis *et al.* 1941; Opsahl 2013). This is equivalent to group-based methods of animal social network construction (Franks *et al.* 2010; Farine & Whitehead 2015). These datasets can be naturally represented as a bipartite network linking individuals to particular groups or aggregations (see *Case Study: Alternative network representations of animal groups*). Frequently, this network is then collapsed into a weighted network using the gambit of the group assumption (see *Case Study: Alternative network representations of animal groups*). Other datasets such as mating networks (Fisher *et al.* 2016; McDonald & Pizzari 2016) or those connecting individuals to locations they have visited (e.g. refuges, watering holes, foraging locations) are also a natural fit to bipartite representations (e.g. (Sah *et al.* 2016)). By representing these networks as bipartite rather than collapsing them to be conventional (unipartite) social networks, key structural information is maintained that would otherwise be lost. This can be very helpful in addressing particular questions. For example, in the case of group-based networks, questions related to choices of group membership for individuals or in breaking down gregariousness into a tendency to be in more and/or larger groups. A wealth of tools is available to analyse bipartite networks, although predominantly tailored to other fields. For example, as well as (generalised) linear model approaches, exponential random graph models have been extended to incorporate bipartite dependency assumptions (Wang *et al.*

2013) and the R package bipartite (Dormann *et al.* 2008) provides capability calculating a range of measures and fitting some more specialist models (albeit with the analysis of ecological rather than social networks in mind). It is important to note that issues with biased sampling (and potentially also identification errors) that apply to dyadic network representations will also apply to the statistical analysis of some bipartite networks, requiring careful consideration during statistical model design.

Tripartite networks

It is also possible to represent animal behaviour with additional layers of complexity. For example, (Manlove *et al.* 2018) demonstrate how it is possible to use a tripartite networks as a conceptual tool to integrate movement and social behaviour. These networks contain three types of node: individuals, spatial locations and time points. Individuals and locations are linked via nodes representing time points. It is then possible to collapse this network into various bipartite and unipartite networks commonly used in behavioural ecology research (see Fig. 2 in (Manlove *et al.* 2018)).

Multilayer networks

Multilayer networks provide a general framework to represent dyadic networks between different types of entity and/or containing different types of social relationships (Kivelä *et al.* 2014). They contain different network layers, with edges possible both within (intra-layer edges) and between (inter-layer edges) them. Because animal populations and groups frequently contain social networks nested within a wider spatial network (Silk *et al.* 2018b; Webber *et al.* 2022), and multiple types of social interaction network can interact with one another within animal groups (Barrett *et al.* 2012; Beisner *et al.* 2015), multilayer networks have great potential as a tool in animal behaviour research (Finn *et al.* 2019).

There are two broad categories of multilayer network. Multiplex networks (typically) contain different sets of interactions between the same set of actors (Kivelä *et al.* 2014). For example, they could be applied in scenarios where a researcher was studying networks of agonistic behaviour, ritualized dominance interactions and submissive behaviours among the same set of individuals (Silk *et al.* 2019). Interconnected networks are most commonly used to represent systems in which layers contain different types of node. This could be different phenotypes (Silk *et al.* 2018c), but could also represent different species or combinations of species and spatial locations. For example, (Silk *et al.* 2018a) used an interconnected network to represent contact networks between wild European badgers *Meles meles*, domestic cattle *Bos taurus* and badger latrine locations. This enabled identification of likely inter- and intraspecific transmission pathways for *Mycobacterium bovis*, the causative agent of bovine tuberculosis.

The use of multilayer approaches has expanded greatly in recent years and there are now a wealth of tools available to analyse them, especially for multiplex networks. These are well summarized by (Finn *et al.* 2019; Finn 2021). muxViz provides a user-friendly interface for basic multilayer network analysis in R (De Domenico *et al.* 2015), with additional R packages tailored to multiplex network analysis including multinet (Magnani *et al.* 2021) and multiplex (Ostoic 2020).

Higher-order network approaches

All of the approaches discussed so far assume that interactions or associations are dyadic – that is that they occur between pairs of individuals. However, it is clear that this represents a simplification for many common types of social data (Musciotto *et al.* 2022). In some cases, explicitly incorporating higher-order interactions (i.e. those simultaneously occurring between more than two individuals) can change our understanding of the emergent properties of a system, for example the spread of pathogens or information (Iacopini *et al.* 2019, 2022; Noonan & Lambiotte 2021). Despite this, explicit higher-order representations of non-human animal social networks have been used only very rarely.

There are three commonly-used ways to represent higher-order interactions in network science: hypergraphs, simplicial sets and simplicial complexes (Silk *et al.* 2022). Hypergraphs are a generalization of dyadic networks that enable (hyper)edges to connect any number of individuals (Battiston *et al.* 2020; Torres *et al.* 2021). For example, three individuals A, B and C observed together in a single group could all be connected with a single hyperedge (or hyperlink). Hypergraph representations of grouping-event based networks is discussed further in the section *Case Study: Alternative network representations of animal groups*.

Simplicial sets represent an alternative mathematical representation of these interactions using set notation (Silk *et al.* 2022). Each simplex in a simplicial set represents either a node/individual (0-simplex), dyadic-interaction (1-simplex) or higher-order interaction (2-simplex, 3-simplex etc.). However, unlike hypergraphs simplicial sets avoid the sub-edge problem (Silk *et al.* 2022); it is possible to represent relationships between individuals in the absence of individuals themselves in simplicial set but not hypergraph representations. For example, imagine a scenario where the presence or outcome of a dominance interaction between individual A and individual D is influenced by an alliance between individuals A, B and C. In this case the 2-simplex (A,B,C) influences the 1-simplex (A,D) even in the absence of B and C. A hypergraph representation cannot capture this component of social structure (Fig. 2).

A simplicial complex is a specific form of simplicial set which must contain all nested lower-order simplices. For example, if a simplicial complex contained the simplex (I,J,K) it

must necessarily also contain the 2-simplices (I,J), (I,K), (J,K), and 1-simplices (I), (J) and (K). In a social context this represents an assumption that any larger interaction or group inevitably includes all possible sub-groups or interactions. While there are cases where this assumption will be met by real-world social behaviour, there are also many cases where not all subsets of individuals within a group can or will have interacted. Therefore, while mathematically convenient, simplicial complexes are likely to be less useful to animal behaviour research than either hypergraphs or simplicial sets.

The power of using higher-order approaches in modelling pathogen spread (Bodó *et al.* 2016; Iacopini *et al.* 2019) and behavioural contagions (Noonan & Lambiotte 2021) has been clearly demonstrated by theoretical work. In behavioural disease ecology higher-order approaches may be particularly useful when there are non-linear dose-response curves (St-Onge *et al.* 2021; Silk *et al.* 2022) and when there is considerable variation in group or aggregation size. For behavioural contagions, higher-order approaches can simplify the representation and modelling of complex contagions (Silk *et al.* 2022). This will be especially powerful when the presence of multiple demonstrators or receivers impacts social learning. However, descriptive measures of higher-order networks will also offer an information rich approach to classifying animal social networks that can extend insights beyond dyadic networks, particularly when studying networks based on co-occurrence in a group (see *Case Study: Alternative network representations of animal groups*). Incorporating hypergraph or simplicial set approaches may even shape how we consider the role of the social environment in indirect genetic effects (Fisher & McAdam 2017; Montiglio *et al.* 2018) if fitness is influenced by higher-order interactions.

Widely accessible implementations of higher-order network approaches are still in their infancy, but there good overviews are provided by (Battiston *et al.* 2020; Torres *et al.* 2021) and guide to available software for visualisation and basic analyses in R, Python and Julia is available in (Silk *et al.* 2022).

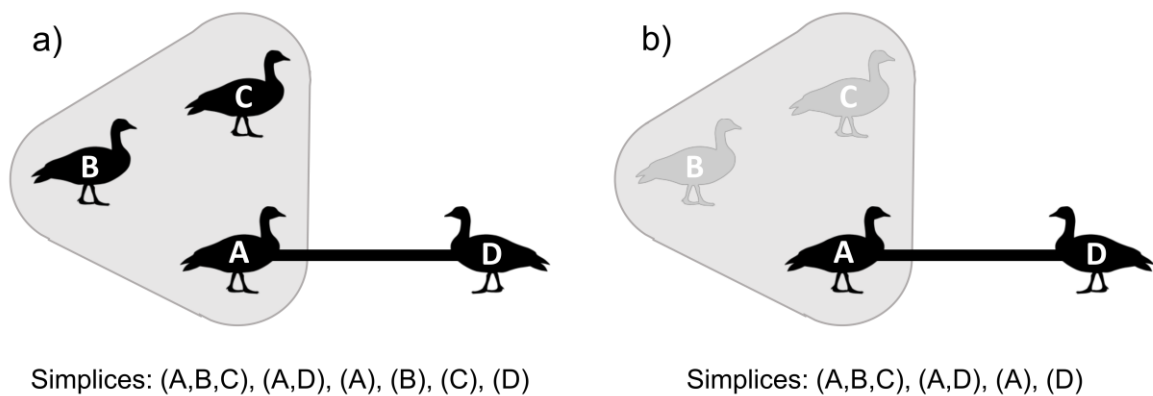


Figure 2. a) Hypergraphs can be used to represent non-dyadic social relationships, associations or interactions (e.g. that between individuals A, B and C) to quantify their role in the structure of animal societies (in this case the interaction between individuals A and D). However, hypergraphs do not naturally represent scenarios when interactions or relationships can have an impact in the absence of the individuals involved. For example, in b) the non-dyadic social relationship between A, B and C can influence the interaction between A and D in the absence of individuals B and C. These scenarios are better captured using simplicial sets, which can include the 2-simplex (A,B,C) without necessarily including the 0-simplices (B) or (C).

Case study: Alternative network representations of animal groups

There is frequently no one single way to represent interactions as a social network. A good example is in considering possible network representations of animal social groups (Fig. 3). Traditionally in animal social network analysis researchers have used the Gambit of the Group assumption (Franks *et al.* 2010) that any pair (dyad) of individuals in a group have associated and therefore share a connection in a weighted social network. The weight of their connection is calculated as a function of the proportion of times they are observed together (with a variety of potential functions available; (Hoppitt & Farine 2018)). However, this representation is a simplification of a more complex network structure which can be captured as either: a) a bipartite network in which one set of nodes represents individuals and another set of nodes observed grouping events or aggregations with edges connecting individuals to the aggregations they occurred in; or b) a [weighted] hypergraph (or simplicial set) in which a hyperedge connects all of the individuals that occurred in each group.

Here we provide a case study (Supplementary Materials) in which three research teams study the same social system. Because they are asking different questions using the dataset they elect to use different network approaches to represent their dataset. Research

Team 1 are interested in whether individual social relationships are assorted by different phenotypic traits. Because the goal of their study is to infer potential social relationships in the population (with the assumption that individuals with strong social relationships will be found together more often in groups) using a dyadic network representation distils relevant information conveniently and provides an effective way to answer their question.

However, in other contexts the weighted (dyadic) network represents a simplification that loses some information about the system. Research Team 2 are more interested in properties of the social groups themselves, specifically whether the social centrality (degree and strength) of individuals is better explained by the size of groups they occur in or the number of times they are in groups (i.e. not observed by themselves) and whether smaller groups tend to consist of subsets of larger groups (i.e. are they nested?). For questions such as these that are more related to the social decision-making of individuals then bipartite network and hypergraph representations are likely to be more suitable as they retain information about group size and composition. In this case, directly modelling the bipartite network could offer real potential, being sufficient to answer their first question and providing ideal tools (nestedness calculations from the ecological networks literature) to answer the second.

The power of higher-order (e.g. hypergraph) approaches becomes valuable if the non-dyadic nature of interactions is likely to be important in some way, especially in studies of contagions, for example. Our third team are disease ecologists who are interested in the potential spread of an emergent pathogen through the population. Because of how it is transmitted they feel that the non-dyadic nature of social interactions in the system may contribute to its spread, and compare hypergraph and dyadic network outbreak models to assess its potential impact. For the set of parameters they consider, the importance of considering the non-dyadic nature of interactions is very evident (although if you change these parameters you will see how this changes according to the size of infectious dose and how this affects the transmission probability of the pathogen).

Collectively, these examples illustrate the advantages of illustrating the same network data in different ways, in particular highlighting how it pays to start with the research question or hypothesis of interest, determine which features of the data are of most importance (and need to be maintained), and then select an appropriate method of representing the network. One thing to note is that the differences between representations can sometimes be conceptual more than practical, for example the group-by-individual matrix used to represent the bipartite network is equivalent to the incidence matrix for a hypergraph representation. It is also important to highlight that it is possible to answer the same question using different representations, for example phenotypic assortativity is apparent in the bipartite network (and analogously the hypergraph) representations as well as the dyadic one. Finally, it may also

be that more some studies combining information from multiple representations is beneficial to addressing research goals.

I have illustrated the example of group-based networks as it is conceptually intuitive, but a similar process is important to follow for other forms of data as well (e.g. association networks based on shared resource use, association networks based on co-capture data).

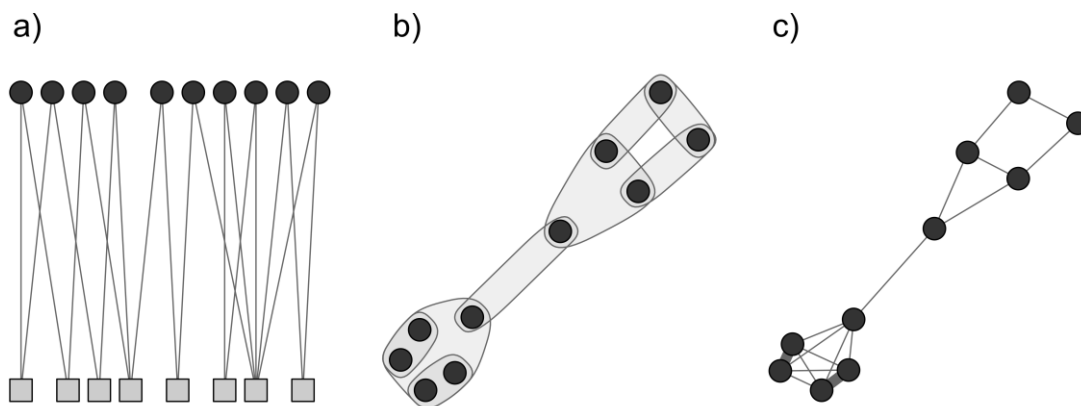


Figure 3. Representations of the same set of animal groups as a) a bipartite network connecting individuals (black circles) to the grouping events (grey squares) they were recorded in; b) a hypergraph connecting individuals (black circles) with hyperedges according to the groups they were observed in; and c) a weighted network in which the weight of dyadic connections represents the number of groups two individuals were observed in together. Each of these approaches retains different levels of information about the social associations that occurred and lends itself naturally to different analytic approaches and research questions.

Dynamic networks

So far all of the network representations have been discussed in a static context, i.e. not considering the fact that interactions or relationships may change over time. However, most network data are dynamic and it can often be helpful to consider how social network structure and the position of individuals within it changes over time (Pinter-Wollman *et al.* 2014; Farine 2017b). For example, when considering pathogen transmission or the spread of information and behaviours, the frequency with which interaction patterns change can be just

as important as social structure in determining the outcome (Evans *et al.* 2020). There are two major ways to consider dynamic network data, as aggregated or snapshot static networks or as time-ordered networks ((Blonder *et al.* 2012); Fig. 4).

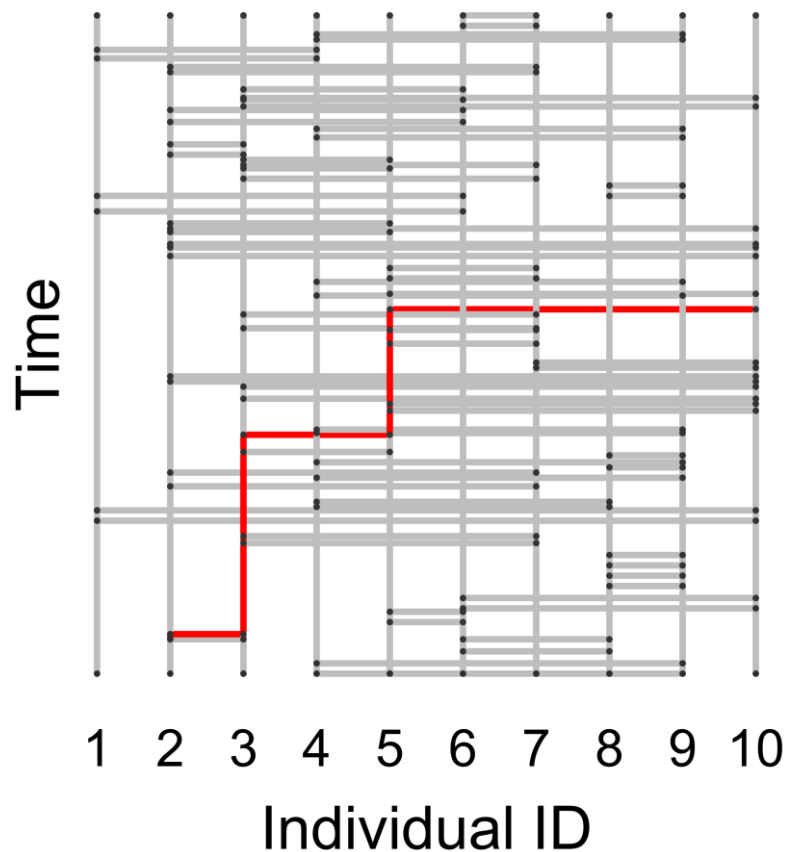


Figure 4. A graphical representation of a time-ordered network using the R package `timeordered`. The start and end of interactions between dyads are indicated by grey horizontal lines and the time at which interactions occurred is indicated by their position on the vertical axis. The red line indicates the shortest possible path between individual 2 and individual 10 while accounting for the order interactions occur in.

The former represents a convenient approach as it facilitates the application of the same (or broadly similar) analytical tools as for static networks. Snapshot networks are a series of networks representing associations or interactions occurring at a series of specific points in time (e.g. the 1-metre proximity networks arising from successive scan samples of a group conducted every hour). For time-aggregated networks there is flexibility in the time periods over which interactions or associations are collated. Often the time periods used are discrete from each other (e.g. daily networks, seasonal networks), however it is also possible

to generate time-aggregated networks based on overlapping windows (e.g. 3-monthly networks that run from January-March, February-April, March-May etc.).

Basic descriptive network measures (whether at the network or individual level) can characterise the stability of overall social structure (Pinter-Wollman *et al.* 2014) and the consistency with which individuals occupy positions within it (Wilson *et al.* 2013). Because time-aggregated networks can also be considered as multiplex networks (with different layers representing different time periods), descriptive measures designed for multilayer networks may also have value for some research questions (Finn *et al.* 2019). Similarly, various statistical models have been designed that specifically analyse these time-aggregated (or snapshot) networks. For example, network autocorrelation models (Silk *et al.* 2017c), exponential random graph models (Lusher *et al.* 2012), and stochastic block models (Matias & Miele 2017) have all been extended to time-aggregated networks. They can be used to answer questions about individual traits, interactions/relationships and network community structure respectively, incorporating temporal variability and time-lagged variables. Further models designed specifically for time-aggregated networks are also available, e.g. stochastic actor-oriented models (Snijders *et al.* 2010; Fisher *et al.* 2017b). It is also possible to apply these statistical models in other ways, for example the epimodel R package (Jenness *et al.* 2018) employs temporal exponential random graph models to facilitate simulations of network epidemiological models. When applying more complex statistical models to time-aggregated networks generated from overlapping time windows (i.e. using subsets of observations that are not fully independent from each other) it is important to ensure this does not violate assumptions of the model being used.

However, while time-aggregated networks are appropriate for answering a wide range of questions in behavioural ecology, they represent a simplification of the true social structure of the data. Consequently, for some applications where the specific order of interaction sequences is important it is necessary to use time-ordered networks. Commonly encountered examples in behavioural ecology are in directed behavioural interaction within-groups such as grooming and dominance interactions. For example, Elo ratings now often used to study dominance hierarchies in animal groups (Neumann *et al.* 2011; Sánchez-Tójar *et al.* 2018; Neumann & Fischer 2022) exploit the information provided by the order interactions occur in. Using time-ordered networks might also be important when answering fine-scale questions about social contagions in networks. For example, retaining time-ordered interactions is important when quantifying social transmission (Silk *et al.* 2017a; Gilbertson *et al.* 2018). Although tools to analyse time-ordered networks are scarcer than for time-aggregated networks, various approaches are available. Measures such as burstiness (Stehlé *et al.* 2010) can quantify the temporal distribution of interactions, which can be important in explaining complex social contagions (Evans *et al.* 2020), for example.

Statistical models include relational event models for directed, time-order networks (Patison *et al.* 2015; Tranmer *et al.* 2015) and dynamic network actor models (Stadtfeld *et al.* 2017) for undirected relationships. The latter could be applied very naturally to study the dynamics of alliances within animal groups, and can be implemented in R using the package *goldfish* (Hollway & Stadtfeld 2022).

Ego networks

Ego networks represent only individuals (nodes) that a single focal individual associates, interacts or shares a social relationship with, along with the social connections between these neighbours (Fig. 5). If desired ego networks can be weighted, signed, multiplex, dynamic etc. in the same way as population or group-level networks (Liu *et al.* 2018; Rezaeipannah *et al.* 2020; Wang *et al.* 2020). The package *egor* (Krenz *et al.* 2022) provides tools for ego network representation, visualisation and analysis in R. While ego networks are commonly used in sociological research, this has not extended to animal behaviour. Ego network analyses and measures typically focus on how embedded an individual is within its network neighbourhood or the extent to which it acts as a bridge between different neighbourhoods (Butts 2008c). While many of these questions can be addressed using measures of individual social network position in group- or population-level networks, ego networks are well-suited to contexts where there are constraints on sampling intensity or wider population data are very sparse. The ego network approach is also well-suited to social network studies that use a focal follow approach (where networks are constructed based on successively following different individuals and recording their social associations and/or interaction partners) as in these cases the full network is constructed from separate ego networks from each focal follow anyway. As a result, while their applications in behavioural ecology will likely be more limited, ego networks are an approach to bear in mind when addressing questions about the social role of particular subsets of individuals especially if there are limitations that require lower levels of sampling. For example, assessing the consistency of an individual's social environment can easily be assessed by calculating the stability of its egonetnetwork. Similarly, the redundancy of an ego within its egonetnetwork provides a good alternative measure of how embedded an individual is within its network. Both these aspects of the social environment have the potential to influence fitness-related traits.

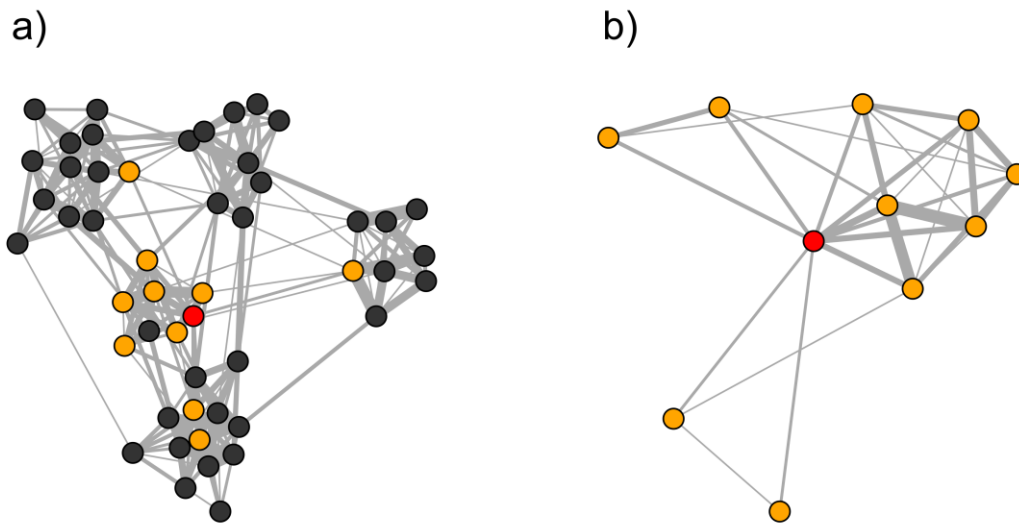


Figure 5. An illustration of a focal individual (red node) within a) its group-level network and b) its ego network. The use of ego network analysis can be used to focus on the immediate social neighbourhood of individuals or when it is prohibitive to sample the entire network.

General recommendations for choosing between approaches

With such diverse approaches available and many applicable to the same datasets, choosing the best representation is not always easy. However, there are a few general rules that can help:

- The research question should always be central to driving the choice of network analyses applied, and this extends beyond how edges are defined (Carter *et al.* 2015; Farine & Whitehead 2015) to broader considerations around network representations. For example, higher-order networks provide valuable additional information when researchers are interested in the ecological or evolutionary effects of larger (non-dyadic) interactions, i.e. non-dyadic effects.
- Sampling constraints are important when selecting representations to use. More complex approaches such as multilayer and dynamic network analyses are more data hungry and require more time-intensive sampling. Ego network analyses can be helpful in situations with patchily distributed sampling effort.
- Collecting “too much” data will only be an issue when making decisions about where to target resources (e.g. ego-networks of individuals in 20 social groups versus a full network of 2 social groups). In other contexts recording data in a way that may allow more complex representations to be used (e.g. recording frequency and duration of

interactions, recording timings of interactions, etc.) will likely be worthwhile even if simpler network representations will probably be used in the long run.

- Some newer approaches have less well-developed or widely accessible analytical tools (e.g. signed and higher-order networks) which may limit the analyses possible without developing functions or algorithms for yourself.
- Some approaches (e.g. weighted multigraphs, higher-order approaches) will be more computationally-intensive than others, and should therefore only be used when they offer a clear advantage for answering a specific research question if computational resources are to be used responsibly.

Conclusions

Network analysis has a huge amount to offer behavioural ecology but frequently represents a major challenge to those encountering it for the first time. Fortunately, there are some valuable introductory papers to offer guidance on data collection, descriptive statistics and statistical modelling in networks (e.g. (Butts 2008c; Pinter-Wollman *et al.* 2014; Farine & Whitehead 2015; Cranmer *et al.* 2016; Farine 2017a; Silk *et al.* 2017c; Hart *et al.* 2021; Hobson *et al.* 2021b)). This overview complements that existing guidance by highlighting the diversity of the applied network analysis toolkit now available and drawing attention to the different ways to think about and represent animal networks. It reveals the power of considering alternative options to quantify networks in animal behaviour research and the potential value of newer approaches to answer key research questions related to social and spatial behaviour.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska Curie grant agreement No. 101023948. I'd like to thank Josh Firth, two anonymous reviewers and the editor for constructive comments on the first submission of the paper. I'd also like to thank Julian Evans and Nina Fefferman for their thoughts on a draft of this paper along with the many other colleagues I have discussed social network analyses with especially Noa-Pinter Wollman, Liz Hobson and David Fisher who have all discussed ideas with me directly related to this paper.

References

- Albery, G.F., Becker, D.J., Brierley, L., Brook, C.E., Christofferson, R.C., Cohen, L.E., *et al.* (2021). The science of the host–virus network. *Nat. Microbiol.*, 6, 1483–1492.
- Allen, J., Weinrich, M., Hoppitt, W. & Rendell, L. (2013). Network-based diffusion analysis reveals cultural transmission of lobtail feeding in humpback whales. *Science (80-.)*, 340, 485–488.
- Aplin, L.M., Farine, D.R., Morand-Ferron, J., Cockburn, A., Thornton, A. & Sheldon, B.C. (2015). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature*, 518, 538–541.
- Balasubramaniam, K.N., Beisner, B.A., Berman, C.M., De Marco, A., Duboscq, J., Koirala, S., *et al.* (2018). The influence of phylogeny, social style, and sociodemographic factors on macaque social network structure. *Am. J. Primatol.*, 80, e22727.
- Barrett, L., Henzi, S.P. & Lusseau, D. (2012). Taking sociality seriously: the structure of multi-dimensional social networks as a source of information for individuals. *Philos. Trans. R. Soc. London B Biol. Sci.*, 367, 2108–2118.
- Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., *et al.* (2020). Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.*, 874, 1–92.
- Beigi, G., Tang, J. & Liu, H. (2016). Signed link analysis in social media networks. In: *Proceedings of the International AAAI Conference on Web and Social Media*. pp. 539–542.
- Beisner, B.A., Jin, J., Fushing, H. & Mccowan, B. (2015). Detection of social group instability among captive rhesus macaques using joint network modeling. *Curr. Zool.*, 61, 70–84.
- Blonder, B., Wey, T.W., Dornhaus, A., James, R. & Sih, A. (2012). Temporal dynamics and network analysis. *Methods Ecol. Evol.*, 3, 958–972.
- Bodó, Á., Katona, G.Y. & Simon, P.L. (2016). SIS epidemic propagation on hypergraphs. *Bull. Math. Biol.*, 78, 713–735.
- Butts, C.C.T. (2008a). Social network analysis with sna. *J. Stat. Softw.*, 24, 13–41.
- Butts, C.T. (2008b). network: a Package for Managing Relational Data in R. *J. Stat. Softw.*, 24, 1–36.
- Butts, C.T. (2008c). Social network analysis: A methodological introduction. *Asian J. Soc. Psychol.*, 11, 13–41.
- Butts, C.T. (2014). Package ‘sna.’
- Butts, C.T. (2015). network: Classes for Relational Data.
- Cantor, M., Maldonado-Chaparro, A.A., Beck, K.B., Brandl, H.B., Carter, G.G., He, P., *et al.* (2021). The importance of individual-to-society feedbacks in animal ecology and evolution. *J. Anim. Ecol.*, 90, 27–44.

- Carter, A.J., Lee, A.E.G. & Marshall, H.H. (2015). Research questions should drive edge definitions in social network studies.
- Connor, R.C., Krützen, M., Allen, S.J., Sherwin, W.B. & King, S.L. (2022). Strategic intergroup alliances increase access to a contested resource in male bottlenose dolphins. *Proc. Natl. Acad. Sci.*, 119, e2121723119.
- Cowl, V.B., Jensen, K., Lea, J., Walker, S.L. & Shultz, S. (2020). Sulawesi crested macaque (*Macaca nigra*) grooming networks are robust to perturbation while individual associations are more labile. *Int. J. Primatol.*, 41, 105–128.
- Cranmer, S.J., Leifeld, P., McClurg, S.D. & Rolfe, M. (2016). Navigating the range of statistical tools for inferential network analysis. *Am. J. Pol. Sci.*
- Croft, D.P., Darden, S.K. & Wey, T.W. (2016). Current directions in animal social networks. *Curr. Opin. Behav. Sci.*, 12, 52–58.
- Croft, D.P., James, R. & Krause, J. (2008). *Exploring animal social networks*. Princeton University Press.
- Croft, D.P., Madden, J.R., Franks, D.W. & James, R. (2011). Hypothesis testing in animal social networks. *Trends Ecol. Evol.*, 26, 502–507.
- Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Syst.*, 1695, 1–9.
- Dakin, R. & Ryder, T.B. (2020). Reciprocity and behavioral heterogeneity govern the stability of social networks. *Proc. Natl. Acad. Sci.*, 117, 2993–2999.
- Davis, A., Gardner, B.B. & Gardner, M.R. (1941). Deep South. The U.
- Dey, C.J. & Quinn, J.S. (2014). Individual attributes and self-organizational processes affect dominance network structure in pukeko. *Behav. Ecol.*, 00, 1–7.
- Ding, Y., Yan, E., Frazho, A. & Caverlee, J. (2009). PageRank for ranking authors in co-citation networks. *J. Am. Soc. Inf. Sci. Technol.*, 60, 2229–2243.
- De Domenico, M., Porter, M.A. & Arenas, A. (2015). MuxViz: a tool for multilayer analysis and visualization of networks. *J. Complex Networks*, 3, 159–176.
- Dormann, C.F., Gruber, B. & Fründ, J. (2008). Introducing the bipartite package: analysing ecological networks. *Interaction*, 1.
- Dupont, Y.L., Trøjelsgaard, K., Hagen, M., Henriksen, M. V, Olesen, J.M., Pedersen, N.M.E., *et al.* (2014). Spatial structure of an individual-based plant–pollinator network. *Oikos*, 123, 1301–1310.
- Edelman, A.J. & McDonald, D.B. (2014). Structure of male cooperation networks at long-tailed manakin leks. *Anim. Behav.*, 97, 125–133.
- Evans, J.C., Silk, M.J., Boogert, N.J. & Hodgson, D.J. (2020). Infected or informed? Social structure and the simultaneous transmission of information and infectious disease. *Oikos*, 129, 1271–1288.

- Facchetti, G., Iacono, G. & Altafini, C. (2011). Computing global structural balance in large-scale signed social networks. *Proc. Natl. Acad. Sci.*, 108, 20953–20958.
- Farine, D.R. (2014). Measuring phenotypic assortment in animal social networks: weighted associations are more robust than binary edges. *Anim. Behav.*, 89, 141–153.
- Farine, D.R. (2017a). A guide to null models for animal social network analysis. *Methods Ecol. Evol.*, 8, 1309–1320.
- Farine, D.R. (2017b). When to choose dynamic versus static social network analysis. *J. Anim. Ecol.*
- Farine, D.R. & Sheldon, B.C. (2015). Selection for territory acquisition is modulated by social network structure in a wild songbird. *J. Evol. Biol.*, 28, 547–556.
- Farine, D.R. & Whitehead, H. (2015). Constructing, conducting and interpreting animal social network analysis. *J. Anim. Ecol.*, 84, 1144–1163.
- Finn, K.R. (2021). Multilayer network analyses as a toolkit for measuring social structure. *Curr. Zool.*, 67, 81–99.
- Finn, K.R., Silk, M.J., Porter, M.A. & Pinter-Wollman, N. (2019). The use of multilayer network analysis in animal behaviour. *Anim. Behav.*, 149, 7–22.
- Fisher, D.N., Ilany, A., Silk, M.J. & Tregenza, T. (2017a). Analysing animal social network dynamics: the potential of stochastic actor-oriented models. *J. Anim. Ecol.*, 86.
- Fisher, D.N., Ilany, A., Silk, M.J. & Tregenza, T. (2017b). Analysing animal social network dynamics: the potential of stochastic actor-oriented models. *J. Anim. Ecol.*, 86, 202–212.
- Fisher, D.N. & McAdam, A.G. (2017). Social traits, social networks and evolutionary biology. *J. Evol. Biol.*, 30, 2088–2103.
- Fisher, D.N., Rodríguez-Muñoz, R. & Tregenza, T. (2016). Comparing pre-and post-copulatory mate competition using social network analysis in wild crickets. *Behav. Ecol.*, 27, 912–919.
- Franks, D.W., Ruxton, G.D. & James, R. (2010). Sampling animal association networks with the gambit of the group. *Behav. Ecol. Sociobiol.*, 64, 493–503.
- Franks, D.W., Weiss, M.N., Silk, M.J., Perryman, R.J.Y. & Croft, D.P. (2021). Calculating effect sizes in animal social network analysis. *Methods Ecol. Evol.*, 12, 33–41.
- Gilbertson, M.L.J., Fountain-Jones, N.M. & Craft, M.E. (2018). Incorporating genomic methods into contact networks to reveal new insights into animal behaviour and infectious disease dynamics. *Behaviour*, 155, 759–791.
- Hamede, R.K., Bashford, J., McCallum, H. & Jones, M. (2009). Contact networks in a wild Tasmanian devil (*Sarcophilus harrisii*) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecol. Lett.*, 12, 1147–1157.

- Hart, J.D.A., Weiss, M.N., Franks, D.W. & Brent, L.J.N. (2021). Bison: A bayesian framework for inference of social networks. *bioRxiv*.
- Hobson, E.A., Mønster, D. & DeDeo, S. (2021a). Aggression heuristics underlie animal dominance hierarchies and provide evidence of group-level social information. *Proc. Natl. Acad. Sci.*, 118, e2022912118.
- Hobson, E.A., Silk, M.J., Fefferman, N.H., Larremore, D.B., Rombach, P., Shai, S., *et al.* (2021b). A guide to choosing and implementing reference models for social network analysis. *Biol. Rev.*, 96, 2716–2734.
- Hollway, J. & Stadtfeld, C. (2022). goldfish: Statistical Network Models for Dynamic Network Data.
- Hoppitt, W.J.E. & Farine, D.R. (2018). Association indices for quantifying social relationships: how to deal with missing observations of individuals or groups. *Anim. Behav.*, 136, 227–238.
- Iacopini, I., Petri, G., Baronchelli, A. & Barrat, A. (2022). Group interactions modulate critical mass dynamics in social convention. *Commun. Phys.*, 5, 1–10.
- Iacopini, I., Petri, G., Barrat, A. & Latora, V. (2019). Simplicial models of social contagion. *Nat. Commun.*, 10, 1–9.
- Ilany, A., Barocas, A., Koren, L., Kam, M. & Geffen, E. (2013). Structural balance in the social networks of a wild mammal. *Anim. Behav.*, 85, 1397–1405.
- Jacoby, D.M.P. & Freeman, R. (2016). Emerging network-based tools in movement ecology. *Trends Ecol. Evol.*, 31, 301–314.
- Jenness, S.M., Goodreau, S.M. & Morris, M. (2018). EpiModel: An R Package for Mathematical Modeling of Infectious Disease over Networks. *J. Stat. Softw.*, 84.
- Jiang, J.Q. (2015). Stochastic block model and exploratory analysis in signed networks. *Phys. Rev. E*, 91, 62805.
- Kirkley, A., Cantwell, G.T. & Newman, M.E.J. (2019). Balance in signed networks. *Phys. Rev. E*, 99, 12320.
- Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y. & Porter, M.A. (2014). Multilayer networks. *J. complex networks*, 2, 203–271.
- Krenz, T., Krivitsky, P.N., Vacca, R., Bojanowski, M. & Herz, A. (2022). egor: Import and Analyse Ego-Centered Network Data.
- Kulahci, I.G., Ghazanfar, A.A. & Rubenstein, D.I. (2018). Knowledgeable lemurs become more central in social networks. *Curr. Biol.*
- Larremore, D.B., Clauset, A. & Jacobs, A.Z. (2014). Efficiently inferring community structure in bipartite networks. *Phys. Rev. E*, 90, 12805.
- Larson, S.M., Ruiz-Lambides, A., Platt, M.L. & Brent, L.J.N. (2018). Social network dynamics precede a mass eviction in group-living rhesus macaques. *Anim. Behav.*, 136, 185–193.

- Liu, C., Wang, J., Zhang, H. & Yin, M. (2018). Mapping the hierarchical structure of the global shipping network by weighted ego network analysis. *Int. J. Shipp. Transp. Logist.*, 10, 63–86.
- Lusher, D., Koskinen, J. & Robins, G. (2012). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press.
- Magnani, M., Rossi, L. & Vega, D. (2021). Analysis of Multiplex Social Networks with R. *J. Stat. Softw.*, 98, 1–30.
- Manlove, K., Aiello, C., Sah, P., Cummins, B., Hudson, P.J. & Cross, P.C. (2018). The ecology of movement and behaviour: a saturated tripartite network for describing animal contacts. *Proc. R. Soc. B Biol. Sci.*, 285, 20180670.
- Matias, C. & Miele, V. (2017). Statistical clustering of temporal networks through a dynamic stochastic block model. *J. R. Stat. Soc. Ser. B (Statistical Methodol.)*, 79, 1119–1141.
- McCowan, B., Beisner, B.A., Capitano, J.P., Jackson, M.E., Cameron, A.N., Seil, S., *et al.* (2011). Network stability is a balancing act of personality, power, and conflict dynamics in rhesus macaque societies. *PLoS One*, 6, e22350.
- McDonald, G.C. & Pizzari, T. (2016). Why patterns of assortative mating are key to study sexual selection and how to measure them. *Behav. Ecol. Sociobiol.*, 70, 209–220.
- Miele, V., Ramos-Jiliberto, R. & Vázquez, D.P. (2020). Core–periphery dynamics in a plant–pollinator network. *J. Anim. Ecol.*, 89, 1670–1677.
- Montiglio, P., McGlothlin, J.W. & Farine, D.R. (2018). Social structure modulates the evolutionary consequences of social plasticity: a social network perspective on interacting phenotypes. *Ecol. Evol.*, 8, 1451–1464.
- Musciotto, F., Papageorgiou, D., Battiston, F. & Farine, D.R. (2022). Beyond the dyad: uncovering higher-order structure within cohesive animal groups. *bioRxiv*.
- Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Irwan, A.M., Agil, M., *et al.* (2011). Assessing dominance hierarchies: validation and advantages of progressive evaluation with Elo-rating. *Anim. Behav.*, 82, 911–921.
- Neumann, C. & Fischer, J. (2022). Extending Bayesian Elo-rating to quantify dominance hierarchy steepness. *bioRxiv*.
- Noonan, J. & Lambiotte, R. (2021). Dynamics of majority rule on hypergraphs. *Phys. Rev. E*, 104, 24316.
- Opsahl, T. (2013). Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Soc. Networks*, 35, 159–167.
- Ostoic, J.A.R. (2020). Algebraic Analysis of Multiple Social Networks with multiplex. *J. Stat. Softw.*, 92, 1–41.
- Patison, K.P., Quintane, E., Swain, D.L., Robins, G. & Pattison, P. (2015). Time is of the essence: an application of a relational event model for animal social networks. *Behav.*

- Ecol. Sociobiol.*, 69, 841–855.
- Pinter-Wollman, N., Hobson, E.A., Smith, J.E., Edelman, A.J., Shizuka, D., de Silva, S., *et al.* (2014). The dynamics of animal social networks: analytical, conceptual, and theoretical advances. *Behav. Ecol.*, 25, 242–255.
- Powell, S.N., Wallen, M.M., Miketa, M.L., Krzyszczyk, E., Foroughirad, V., Bansal, S., *et al.* (2020). Sociality and tattoo skin disease among bottlenose dolphins in Shark Bay, Australia. *Behav. Ecol.*, 31, 459–466.
- Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). Network thinking in ecology and evolution. *Trends Ecol. Evol.*, 20, 345–353.
- Rezaeipanah, A., Ahmadi, G. & Sechin Matoori, S. (2020). A classification approach to link prediction in multiplex online ego-social networks. *Soc. Netw. Anal. Min.*, 10, 1–16.
- Ross, C.T., McElreath, R. & Redhead, D. (2022). Modelling human and non-human animal network data in R using STRAND. *bioRxiv*.
- Sah, P., Nussear, K.E., Esque, T.C., Aiello, C.M., Hudson, P.J. & Bansal, S. (2016). Inferring social structure and its drivers from refuge use in the desert tortoise, a relatively solitary species. *Behav. Ecol. Sociobiol.*, 70, 1277–1289.
- Sánchez-Tójar, A., Schroeder, J. & Farine, D.R. (2018). A practical guide for inferring reliable dominance hierarchies and estimating their uncertainty. *J. Anim. Ecol.*, 87, 594–608.
- Schoch, D. (2020). signnet: An R package to analyze signed networks.
- Shizuka, D. & McDonald, D.B. (2012). A social network perspective on measurements of dominance hierarchies. *Anim. Behav.*, 83, 925–934.
- Silk, M.J., Cant, M.A., Cafazzo, S., Natoli, E. & McDonald, R.A. (2019). Elevated aggression is associated with uncertainty in a network of dog dominance interactions. *Proc. R. Soc. B Biol. Sci.*, 286.
- Silk, M.J., Croft, D.P., Delahay, R.J., Hodgson, D.J., Boots, M., Weber, N., *et al.* (2017a). Using social network measures in wildlife disease ecology, epidemiology, and management. *Bioscience*, 67.
- Silk, M.J., Croft, D.P., Delahay, R.J., Hodgson, D.J., Boots, M., Weber, N., *et al.* (2017b). Using Social Network Measures in Wildlife Disease Ecology, Epidemiology, and Management. *Bioscience*, 67, 245–257.
- Silk, M.J., Croft, D.P., Delahay, R.J., Hodgson, D.J., Weber, N., Boots, M., *et al.* (2017c). The application of statistical network models in disease research. *Methods Ecol. Evol.*
- Silk, M.J., Drewe, J.A., Delahay, R.J., Weber, N., Steward, L.C., Wilson-Aggarwal, J., *et al.* (2018a). Quantifying direct and indirect contacts for the potential transmission of infection between species using a multilayer contact network.
- Silk, M.J., Finn, K.R., Porter, M.A. & Pinter-Wollman, N. (2018b). Can Multilayer Networks Advance Animal Behavior Research? *Trends Ecol. Evol.*

- Silk, M.J., Weber, N.L., Steward, L.C., Hodgson, D.J., Boots, M., Croft, D.P., *et al.* (2018c). Contact networks structured by sex underpin sex-specific epidemiology of infection. *Ecol. Lett.*, 21.
- Silk, M.J., Wilber, M.Q. & Fefferman, N.H. (2022). Capturing complex interactions in disease ecology with simplicial sets. *Ecol. Lett.*, 25, 2217–2231.
- Snijders, T.A.B., Van de Bunt, G.G. & Steglich, C.E.G. (2010). Introduction to stochastic actor-based models for network dynamics. *Soc. Networks*, 32, 44–60.
- St-Onge, G., Sun, H., Allard, A., Hébert-Dufresne, L. & Bianconi, G. (2021). Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks. *Phys. Rev. Lett.*, 127, 158301.
- Stadtfeld, C., Hollway, J. & Block, P. (2017). Dynamic network actor models: Investigating coordination ties through time. *Sociol. Methodol.*, 47, 1–40.
- Stehlé, J., Barrat, A. & Bianconi, G. (2010). Dynamical and bursty interactions in social networks. *Phys. Rev. E*, 81, 35101.
- Strickland, K., Levensgood, A., Foroughirad, V., Mann, J., Krzyszczuk, E. & Frère, C.H. (2017). A framework for the identification of long-term social avoidance in longitudinal datasets. *R. Soc. open Sci.*, 4, 170641.
- Torres, L., Blevins, A.S., Bassett, D. & Eliassi-Rad, T. (2021). The why, how, and when of representations for complex systems. *SIAM Rev.*, 63, 435–485.
- Tranmer, M., Marcum, C.S., Morton, F.B., Croft, D.P. & de Kort, S.R. (2015). Using the relational event model (REM) to investigate the temporal dynamics of animal social networks. *Anim. Behav.*, 101, 99–105.
- Valverde, S., Vidiella, B., Montañez, R., Fraile, A., Sacristán, S. & García-Arenal, F. (2020). Coexistence of nestedness and modularity in host–pathogen infection networks. *Nat. Ecol. Evol.*, 4, 568–577.
- VanderWaal, K.L., Atwill, E.R., Isbell, L. & McCowan, B. (2014). Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). *J. Anim. Ecol.*, 83, 406–414.
- Wang, J., Yang, N. & Guo, M. (2020). Ego-network stability and exploratory innovation: the moderating role of knowledge networks. *Manag. Decis.*
- Wang, P., Pattison, P. & Robins, G. (2013). Exponential random graph model specifications for bipartite networks—A dependence hierarchy. *Soc. Networks*, 35, 211–222.
- Webber, Q., Albery, G., Farine, D.R., Pinter-Wollman, N., Sharma, N., Spiegel, O., *et al.* (2022). Behavioural ecology at the spatial-social interface.
- Wey, T., Blumstein, D.T., Shen, W. & Jordán, F. (2008). Social network analysis of animal behaviour: a promising tool for the study of sociality. *Anim. Behav.*, 75, 333–344.
- White, L.A., Forester, J.D. & Craft, M.E. (2017). Using contact networks to explore

mechanisms of parasite transmission in wildlife. *Biol. Rev.*, 92, 389–409.

Wice, E.W. & Saltz, J.B. (2021). Selection on heritable social network positions is context-dependent in *Drosophila melanogaster*. *Nat. Commun.*, 12, 1–9.

Wilson, A.D.M., Krause, S., Dingemanse, N.J. & Krause, J. (2013). Network position: a key component in the characterization of social personality types. *Behav. Ecol. Sociobiol.*, 67, 163–173.

Young, J.-G., Cantwell, G.T. & Newman, M.E.J. (2020). Bayesian inference of network structure from unreliable data. *J. Complex Networks*, 8, cnaa046.

Network Representations Case Study

Matthew Silk

In this case study I will use association-based (gambit of the group) data to illustrate the value of using different network representations to ask different research questions.

I will first simulate a data set of observations of individuals in groups, in which group membership is determined by an underlying social network that depends on various phenotypic traits.

I will then set up three *research teams*, each with different questions, that will analyse this data set in different ways.

Team 1 will use weighted **dyadic networks** to test whether individual social relationships are assorted by different phenotypic traits

Team 2 will use **bipartite networks** to: a) quantify whether differences in individual gregariousness are better explained by the number of groups they occur in or the average size of groups that they occur in; and b) ask if different social groups are nested (i.e. smaller social groups are typically subsets of larger groups).

Team 3 are interested in modelling the transmission of an emerging viral disease with aerosol transmission. They are interested in the importance of taking into account nondyadic (higher-order) interactions so compare insights from a **hypergraph** and dyadic network modelling approach.

The first step is to set a seed for the random number generator

```
set.seed(6)
```

The next step is to load the R packages we will need for the case study (it may be that you need to install some of these packages first)

```
library(igraph)
library(boot)
library(genNetDem)
library(asnipe)
library(assortnet)
library(bipartite)
library(vegan)
library(ergm)
library(latentnet)
library(sbm)
```

The next step is to provide additional functions that are needed to simulate the data. The functions here are adapted from the genNetDem R package (<https://github.com/NETDEM-project/genNetDem>). I will not show this code chunk in the output.

I can then simulate the data set.

First, the individuals and their traits.

```
#Population size
n<-50

#Individual identities
id<-seq(1,50,1)

#Trait 1 - sex (female or male)
sex<-sample(c("M","F"),50,replace=TRUE)

#Trait 2 - size (normally distributed)
size<-rnorm(50,0,1)

#Trait 3 - colour (3-level categorical)
colour<-sample(c("blue","orange","red"),50,replace=TRUE)
```

Second, the underlying social network. I will simulate an underlying network in which there is assortment by colour and females have more connections than males.

```

#Empty adjacency matrix
adj_mat<-matrix(0,nr=n,nc=n)

#Effect sizes for edge probabilities
b_prob<-0.05 #baseline edge probability
c_eff<-1.75 #effect size of being same colour on edge probability (logit scale)
s_eff<-0.5 #effect size of being female on edge probability (logit scale)

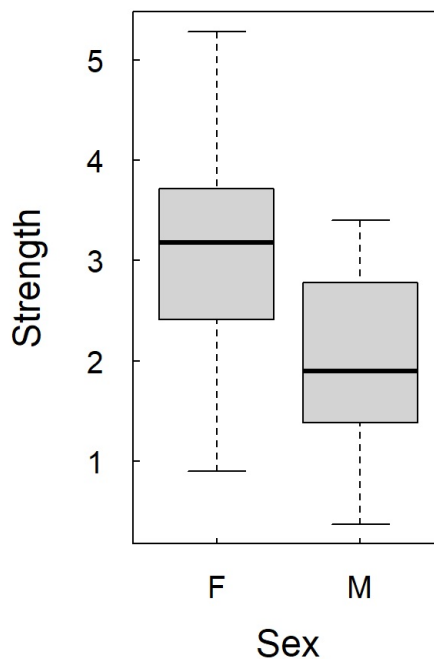
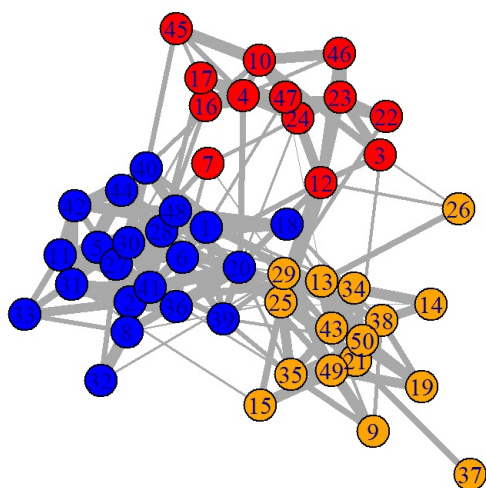
#Generate social connections (initially binary)
for(i in 1:(n-1)){
  for(j in (i+1):n){
    t_prob<-boot::inv.logit(boot::logit(b_prob)+
                          (colour[i]==colour[j])*c_eff+
                          ((sex[i]=="F")*s_eff+(sex[j]=="F")*s_eff)/2)
    adj_mat[i,j]<-adj_mat[j,i]<-rbinom(1,1,t_prob)
  }
}

#Effect sizes for edge weights
b_weight<-0.2 #baseline edge weight
b_sd<-0.5 #variation in edge weight
c_eff2<-1 #effect size of being same colour on edge weight (logit scale)
s_eff2<-0.25 #effect size of being female on edge weight (logit scale)

for(i in 1:(n-1)){
  for(j in (i+1):n){
    t_weight<-boot::inv.logit(rnorm(1,
                                   boot::logit(b_weight)+
                                   (colour[i]==colour[j])*c_eff2+
                                   ((sex[i]=="F")*s_eff2+(sex[j]=="F")*s_eff2)/2,
                                   b_sd))
    adj_mat[i,j]<-adj_mat[j,i]<-adj_mat[i,j]*t_weight
  }
}

#Plot to check that the simulation is generating expected results
net<-igraph::graph.adjacency(adj_mat,mode="undirected",weighted=TRUE)
par(mfrow=c(1,2),tcl=0.2,las=1)
par(mar=c(0,0,0,0))
plot(net,vertex.color=colour,edge.width=(E(net)$weight*10)^1.2)
par(mar=c(5,5,2,2))
boxplot(igraph::strength(net)~sex,xlab="Sex",ylab="Strength",las=1,cex.axis=1.25,cex.lab=1.5)

```



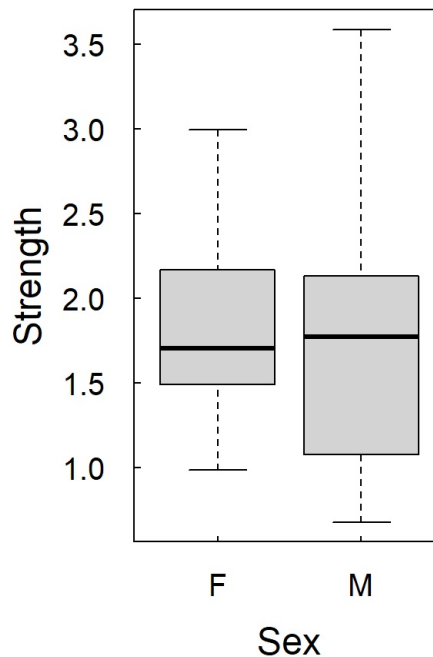
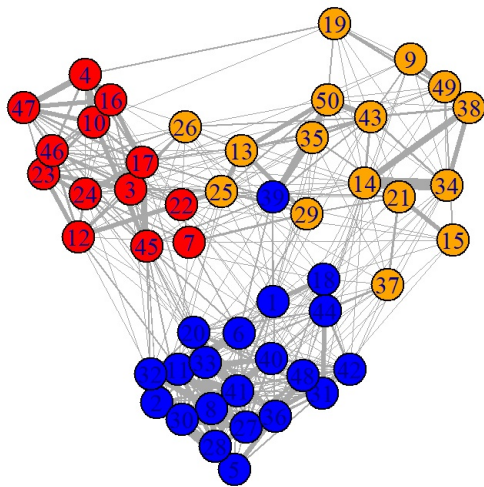
Third, the groups. To start with we use the (adapted) functionality of genNetDem to simulate groups with a mean size of two over 10 timesteps. Every individual will occur at each timestep but will sometimes be in a group of one (alone).


```
#Simulate the groups
groups<-interaction_generation_simul(id=id,adj_mat=adj_mat,mean_group_size=2,
                                   n_ts=10,float=0.000000001,pm=50,pow=4)
```

```
#Check that the group-based network is similar to the underlying network and that the
#effect of sex is maintained
int_mat<-asnipe::get_network(groups[[1]])
```

```
## Generating 50 x 50 matrix
```

```
net<-igraph::graph.adjacency(int_mat,mode="undirected",weighted=TRUE)
par(mfrow=c(1,2))
par(mar=c(0,0,0,0),tcl=0.2,las=1)
plot(net,vertex.color=colour,edge.width=(E(net)$weight*10)^1.2)
par(mar=c(5,5,2,2))
boxplot(igraph::strength(net)~sex,xlab="Sex",ylab="Strength",las=1,cex.axis=1.25,cex.lab=1.5)
```



```
#Use netlm to inspect the correlation between the two adjacency matrices
sna::netlm(int_mat,adj_mat,diag=FALSE,nullhyp="qapspp")
```

```
##
## OLS Network Model
##
## Coefficients:
##           Estimate Pr(<=b) Pr(>=b) Pr(>=|b|)
## (intercept) 0.01905247 0.984 0.016 0.016
## x1          0.34118332 1.000 0.000 0.000
##
## Residual standard error: 0.05531 on 2448 degrees of freedom
## F-statistic: 1886 on 1 and 2448 degrees of freedom, p-value: 0
## Multiple R-squared: 0.4352 Adjusted R-squared: 0.4349
```

Note that the correlation between the two networks is relatively good (it would be stronger if we sampled groups for longer) and that assortment is still apparent (by eye at least!), but we lose the weaker sex effect, or at least the medians become much more similar.

Now we have our data set we can follow the three teams' investigations to see the advantages of using different representations.

Our three teams have access to the same set of observations:

- a group-by-individual matrix of the observed grouping events over 10 time steps (note we are making the unrealistic assumption that they managed to observe every group).
- individual-level data on sex, size and colour.

As a reminder **team 1** are interested in whether individual social relationships are assorted by different phenotypic traits.

When answering questions related to underlying social relationships within the population, collapsing the bipartite network of groups and individuals to a dyadic network is typically helpful.

This network could be a weighted network that uses association indices or potentially a multigraph in which each edge within a dyad represents an observation of them together in a group. Here they take the former approach using the R package `asnipe` (<https://cran.r-project.org/web/packages/asnipe/asnipe.pdf>) for convenience. They take a simple approach to measuring assortment by colour using the R package `assortnet` (<https://cran.r-project.org/web/packages/assortnet/assortnet.pdf>).

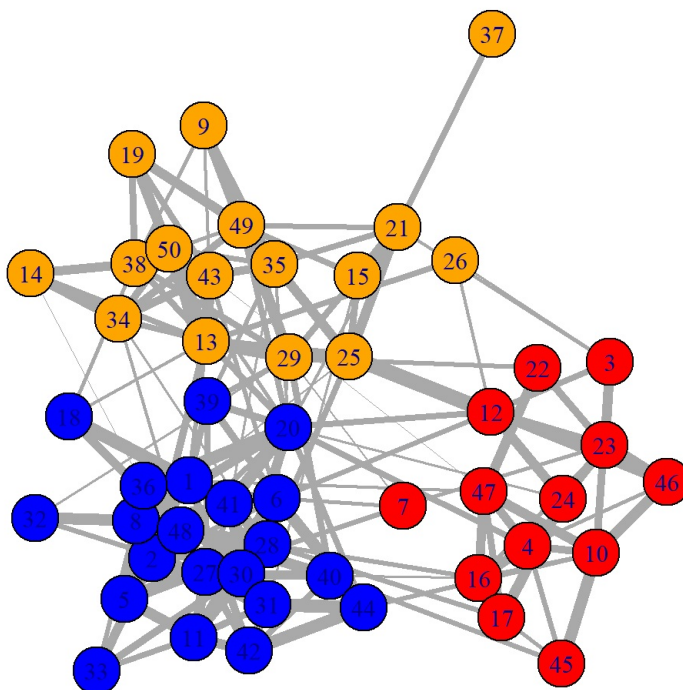
```
#They first generate the projected, weighted network using the simple ratio index for edge weightings  
int_mat<-asnipe::get_network(groups[[1]])
```

```
## Generating 50 x 50 matrix
```

```
#They can formally test for assortment using the assortnet package  
assortnet::assortment.discrete(int_mat,types=colour,weighted=TRUE,SE=TRUE)
```

```
## $r  
## [1] 0.7016485  
##  
## $se  
## [1] 0.02224101  
##  
## $mixing_matrix  
##           blue      red      orange      ai  
## blue  0.46095865 0.02778514 0.03560836 0.5243522  
## red   0.02778514 0.19317145 0.02785625 0.2488128  
## orange 0.03560836 0.02785625 0.16337040 0.2268350  
## bi    0.52435216 0.24881284 0.22683501 1.0000000
```

```
#They can then plot the projected network to look for assortment visually  
net<-igraph::graph.adjacency(adj_mat,mode="undirected",weighted=TRUE)  
par(mar=c(0,0,0,0))  
plot(net,vertex.color=colour,edge.width=(E(net)$weight*10)^1.2)
```



```
#As I have simulated our data I can also test what the assortment by colour would be in  
#the underlying network  
#It would also be possible to repeat this test with more or fewer timesteps sampled  
assortnet::assortment.discrete(adj_mat,types=colour,weighted=TRUE,SE=TRUE)
```

```
## $r
## [1] 0.7448817
##
## $se
## [1] 0.02912628
##
## $mixing_matrix
##           blue      red      orange      ai
## blue  0.43058245 0.02435121 0.04168590 0.4966196
## red   0.02435121 0.16635611 0.01332123 0.2040285
## orange 0.04168590 0.01332123 0.24434477 0.2993519
## bi    0.49661956 0.20402855 0.29935189 1.0000000
```

As an example, I have gone as far as is necessary to show that the dyadic representation is an effective tool to test these types of question for association-based data. For more formal analyses homophily/assortment by phenotypic traits can suit statistical inference using well-designed permutation tests (<https://doi.org/10.1111/brv.12775>) (e.g. node swaps in this case with perfect sampling). However, especially with less even sampling, then an alternative approach would be to use dyadic regressions. Given that only a small number of grouping events have been sampled, a method that incorporates error in network edge weights such as the BISO framework (<https://doi.org/10.1101/2021.12.20.473541>) would be ideal.

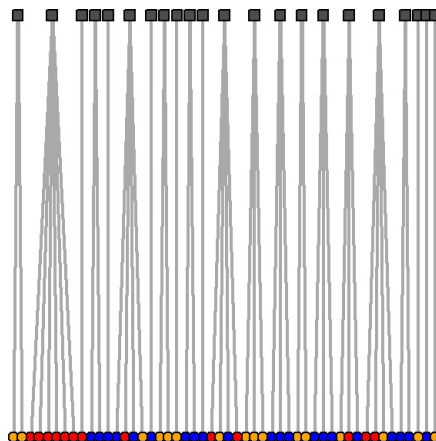
As a reminder **team 2** are more interested in questions related to groups, both what aspects of group decision making lead individuals to be more spontaneous, and whether smaller groups tend to be subsets of larger groups in this fission-fusion society.

For questions directly related to groups themselves then retaining the **bipartite network** is often helpful. The reasons for this become clear when they plot one sampling window of the bipartite network.

```
#Extract the group-by-individual matrix for just the first timestep
gbi_start<-groups[[1]][groups[[2]]<2,]

#Generate a bipartite network in igraph
bipart<-igraph::graph_from_incidence_matrix(gbi_start,directed=FALSE)
layers<-V(bipart)$type+1

#Plot network
plot(bipart,layout=layout.bipartite(bipart),vertex.label=NA,vertex.size=5,vertex.shape=ifelse(layers==1,"square",
"circle"),edge.width=2,vertex.color=c(rep("gray30",sum(layers==1)),colour))
```



Using the bipartite projection it is easy to tease apart how a tendency to be found in large groups versus small groups impacts an individual's social network position. With the full bipartite network we can examine how individuals are linked through groups and learn more about how groups are linked (indirectly) with each other.

With this in mind they plot the full bipartite network (note the different format here)

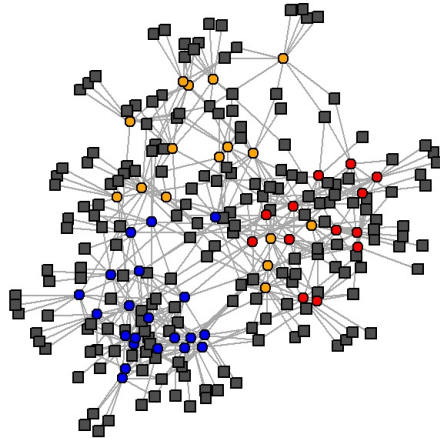
```

#Extract full group-by-individual matrix
gbi_full<-groups[[1]]

#Generate bipartite network in igraph
bipart<-igraph::graph_from_incidence_matrix(gbi_full,directed=FALSE)
layers<-V(bipart)$type+1

#Plot network
plot(bipart,vertex.label=NA,vertex.size=5,vertex.shape=ifelse(layers==1,"square","circle"),edge.width=1,vertex.col=rep("gray30",sum(layers==1)),colour)

```



Team 2 starts with a simple analysis, using the incidence matrix of the bipartite network to calculate the number of groups an individual occurs in (i.e. number of times they were not alone in a given timestep) and the mean size of groups that an individual occurs in. Their aim here is to see whether differences in gregariousness (both strength and degree from the projected dyadic network) are best explained by how many groups an individual is in or the size of these groups.

```

#Extract group-by-individual matrix without lone individuals
gbi_full2<-gbi_full[rowSums(gbi_full)>1,]

#Record group sizes from bipartite network for each individual
group_sizes<-list()
for(i in 1:ncol(gbi_full2)){
  group_sizes[[i]]<-rowSums(gbi_full2[gbi_full2[,i]==1,])
}

#Calculate the number of groups (of more than one) and mean group size for each individual
gn<-sapply(group_sizes,length)
gs<-sapply(group_sizes,mean)

#Calculate the strength and degree for each individual in the dyadic network
str1<-colSums(int_mat)
deg1<-colSums(sign(int_mat))

#Scale the number of groups and mean group size to both have a mean of zero and unit variance
sc_gn<-scale(gn)
sc_gs<-scale(gs)

#Fit models for both strength (Gaussian linear model) and degree (Poisson generalised linear model) with both the
number of groups and mean group size as explanatory variables
str_mod<-lm(str1~sc_gn+sc_gs)
deg_mod<-glm(deg1~sc_gn+sc_gs,family="poisson")

```

They look at the model summary to see that mean group size is more important (approximately 4x as important by comparing the effect sizes from the standardized variables) than number of groups in explaining an individual's degree.

```
summary(deg_mod)
```

```
##
## Call:
## glm(formula = deg1 ~ sc_gn + sc_gs, family = "poisson")
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6163 -0.5708 -0.1645  0.5387  1.9634
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.75788    0.03591  76.808 < 2e-16 ***
## sc_gn        0.04921    0.03624   1.358  0.174
## sc_gs        0.17409    0.03465   5.024 5.06e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 59.912  on 49  degrees of freedom
## Residual deviance: 31.922  on 47  degrees of freedom
## AIC: 267.16
##
## Number of Fisher Scoring iterations: 4
```

They find the same thing for strength, but now only (approximately) twice as important.

```
summary(str_mod)
```

```
##
## Call:
## lm(formula = str1 ~ sc_gn + sc_gs)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.271323 -0.070420  0.005911  0.048769  0.285382
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.81426    0.01541  117.76 <2e-16 ***
## sc_gn        0.26476    0.01566   16.91 <2e-16 ***
## sc_gs        0.54558    0.01566   34.85 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1089 on 47 degrees of freedom
## Multiple R-squared:  0.9723, Adjusted R-squared:  0.9711
## F-statistic: 824.5 on 2 and 47 DF,  p-value: < 2.2e-16
```

This is not the only question they are interested in related to groups. They also suspect that due to the tendency for individuals to assort by colour (they have been speaking to team 1) that the bipartite network might be nested so that small groups tend to be subsets of larger groups. Given this could have important ecological and evolutionary consequences they are keen to test this formally.

Fortunately, it is easy to directly calculate the nestedness of a bipartite network. Here we calculate the nestedness temperature using the `vegan` package, where lower values indicate a more nested network.

```
nn<-vegan::nestedtemp(gbi_full2)
```

Unfortunately, this value does not mean very much on its own. Therefore they decide they need to use a permutation approach to see whether the nestedness value they calculate for the observed network is higher than you would expect if groups were formed at random.

Given they are fortunate and have no major sampling limitations or spatial constraints on interactions they choose a simple datastream permutation in which individuals are swapped between groups observed on the same day/at the same time.

They use a Markov chain approach to generate their reference distribution using a burn-in of 500 swaps (in which no samples are taken) and a thinning interval of 10 (saving the bipartite network every 10 swaps after this). They use three separate Markov chains to improve the reliability of their analysis. They therefore generate a reference distribution of 3000 permuted networks (the R objects `rgbisA`, `rgbisB` and `rgbisC`, which are then combined into the overall `rgbis` list).

They then calculate the nestedness of all of the bipartite networks in their reference distribution and plot the histogram of this reference distribution (with each colour representing a different chain).

```
##First Markov Chain
```

```
gbi_t<-gbi_full2
rgbis<-list()
```

```
day<-groups[[2]]
```

```
#The burn-in period for the first Markov Chain
```

```
for(i in 1:500){  
  #sample an individual/grouping-event  
  pind<-which(gbi_t>0,arr.ind=TRUE)  
  tind1<-pind[sample(1:nrow(pind),1),]  
  #record the day on which that individual/grouping-event occurred  
  td<-which(day==day[tind1[1]])  
  #sample a second individual/grouping-event that occurs on the same day  
  pind2<-pind[which(pind[,1]%in%td),]  
  tind2<-pind2[sample(1:nrow(pind2),1),]  
  #If additional constraints are met then conduct swap  
  if(tind1[1]!=tind2[1]&tind1[2]!=tind2[2]){  
    if(gbi_t[tind1[1],tind2[2]]==0&gbi_t[tind2[1],tind1[2]]==0){  
      gbi_t2<-gbi_t  
      gbi_t2[tind2[1],tind1[2]]<-gbi_t[tind1[1],tind1[2]]  
      gbi_t2[tind1[1],tind1[2]]<-gbi_t[tind2[1],tind1[2]]  
      gbi_t2[tind1[1],tind2[2]]<-gbi_t[tind2[1],tind2[2]]  
      gbi_t2[tind2[1],tind2[2]]<-gbi_t[tind1[1],tind2[2]]  
      gbi_t<-gbi_t2  
    }  
  }  
}
```

```
#The sampling period for the first Markov Chain
```

```
c<-1  
for(i in 1:10000){  
  pind<-which(gbi_t>0,arr.ind=TRUE)  
  tind1<-pind[sample(1:nrow(pind),1),]  
  td<-which(day==day[tind1[1]])  
  pind2<-pind[which(pind[,1]%in%td),]  
  tind2<-pind2[sample(1:nrow(pind2),1),]  
  if(tind1[1]!=tind2[1]&tind1[2]!=tind2[2]){  
    if(gbi_t[tind1[1],tind2[2]]==0&gbi_t[tind2[1],tind1[2]]==0){  
      gbi_t2<-gbi_t  
      gbi_t2[tind2[1],tind1[2]]<-gbi_t[tind1[1],tind1[2]]  
      gbi_t2[tind1[1],tind1[2]]<-gbi_t[tind2[1],tind1[2]]  
      gbi_t2[tind1[1],tind2[2]]<-gbi_t[tind2[1],tind2[2]]  
      gbi_t2[tind2[1],tind2[2]]<-gbi_t[tind1[1],tind2[2]]  
      gbi_t<-gbi_t2  
    }  
  }  
  #This is where we save the swaps. Notice we only save every 10th swap  
  if(i%10==0){  
    rgbis[[c]]<-gbi_t  
    c<-c+1  
  }  
}  
rgbisA<-rgbis
```

```
####
```

```
##Second Markov Chain
```

```
gbi_t<-gbi_full2  
rgbis<-list()  
day<-groups[[2]]
```

```
#Burn-in
```

```
for(i in 1:500){  
  #sample an individual/grouping-event  
  pind<-which(gbi_t>0,arr.ind=TRUE)  
  tind1<-pind[sample(1:nrow(pind),1),]  
  #record the day on which that individual/grouping-event occurred  
  td<-which(day==day[tind1[1]])  
  #sample a second individual/grouping-event that occurs on the same day  
  pind2<-pind[which(pind[,1]%in%td),]  
  tind2<-pind2[sample(1:nrow(pind2),1),]  
  #If additional constraints are met then conduct swap  
  if(tind1[1]!=tind2[1]&tind1[2]!=tind2[2]){  
    if(gbi_t[tind1[1],tind2[2]]==0&gbi_t[tind2[1],tind1[2]]==0){  
      gbi_t2<-gbi_t  
      gbi_t2[tind2[1],tind1[2]]<-gbi_t[tind1[1],tind1[2]]  
      gbi_t2[tind1[1],tind1[2]]<-gbi_t[tind2[1],tind1[2]]  
      gbi_t2[tind1[1],tind2[2]]<-gbi_t[tind2[1],tind2[2]]  
      gbi_t2[tind2[1],tind2[2]]<-gbi_t[tind1[1],tind2[2]]  
      gbi_t<-gbi_t2  
    }  
  }  
}
```

```

}
}

#Sampling
c<-1
for(i in 1:10000){
  pind<-which(gbi_t>0,arr.ind=TRUE)
  tind1<-pind[sample(1:nrow(pind),1),]
  td<-which(day==day[tind1[1]])
  pind2<-pind[which(pind[,1]%in%td),]
  tind2<-pind2[sample(1:nrow(pind2),1),]
  if(tind1[1]!=tind2[1]&tind1[2]!=tind2[2]){
    if(gbi_t[tind1[1],tind2[2]]==0&gbi_t[tind2[1],tind1[2]]==0){
      gbi_t2<-gbi_t
      gbi_t2[tind2[1],tind1[2]]<-gbi_t[tind1[1],tind1[2]]
      gbi_t2[tind1[1],tind1[2]]<-gbi_t[tind2[1],tind1[2]]
      gbi_t2[tind1[1],tind2[2]]<-gbi_t[tind2[1],tind2[2]]
      gbi_t2[tind2[1],tind2[2]]<-gbi_t[tind1[1],tind2[2]]
      gbi_t<-gbi_t2
    }
  }
  #This is where we save the swaps. Notice we only save every 10th swap
  if(i%10==0){
    rgbis[[c]]<-gbi_t
    c<-c+1
  }
}
rgbisB<-rgbis

####

##Third Markov Chain

gbi_t<-gbi_full2
rgbis<-list()
day<-groups[[2]]

#Burn-in
for(i in 1:500){
  #sample an individual/grouping-event
  pind<-which(gbi_t>0,arr.ind=TRUE)
  tind1<-pind[sample(1:nrow(pind),1),]
  #record the day on which that individual/grouping-event occurred
  td<-which(day==day[tind1[1]])
  #sample a second individual/grouping-event that occurs on the same day
  pind2<-pind[which(pind[,1]%in%td),]
  tind2<-pind2[sample(1:nrow(pind2),1),]
  #If additional constraints are met then conduct swap
  if(tind1[1]!=tind2[1]&tind1[2]!=tind2[2]){
    if(gbi_t[tind1[1],tind2[2]]==0&gbi_t[tind2[1],tind1[2]]==0){
      gbi_t2<-gbi_t
      gbi_t2[tind2[1],tind1[2]]<-gbi_t[tind1[1],tind1[2]]
      gbi_t2[tind1[1],tind1[2]]<-gbi_t[tind2[1],tind1[2]]
      gbi_t2[tind1[1],tind2[2]]<-gbi_t[tind2[1],tind2[2]]
      gbi_t2[tind2[1],tind2[2]]<-gbi_t[tind1[1],tind2[2]]
      gbi_t<-gbi_t2
    }
  }
}

#Sampling
c<-1
for(i in 1:10000){
  pind<-which(gbi_t>0,arr.ind=TRUE)
  tind1<-pind[sample(1:nrow(pind),1),]
  td<-which(day==day[tind1[1]])
  pind2<-pind[which(pind[,1]%in%td),]
  tind2<-pind2[sample(1:nrow(pind2),1),]
  if(tind1[1]!=tind2[1]&tind1[2]!=tind2[2]){
    if(gbi_t[tind1[1],tind2[2]]==0&gbi_t[tind2[1],tind1[2]]==0){
      gbi_t2<-gbi_t
      gbi_t2[tind2[1],tind1[2]]<-gbi_t[tind1[1],tind1[2]]
      gbi_t2[tind1[1],tind1[2]]<-gbi_t[tind2[1],tind1[2]]
      gbi_t2[tind1[1],tind2[2]]<-gbi_t[tind2[1],tind2[2]]
      gbi_t2[tind2[1],tind2[2]]<-gbi_t[tind1[1],tind2[2]]
      gbi_t<-gbi_t2
    }
  }
  #This is where we save the swaps. Notice we only save every 10th swap

```

```

if(i%%10==0){
  rgbis[[c]]<-gbi_t
  c<-c+1
}
}
rgbisC<-rgbis

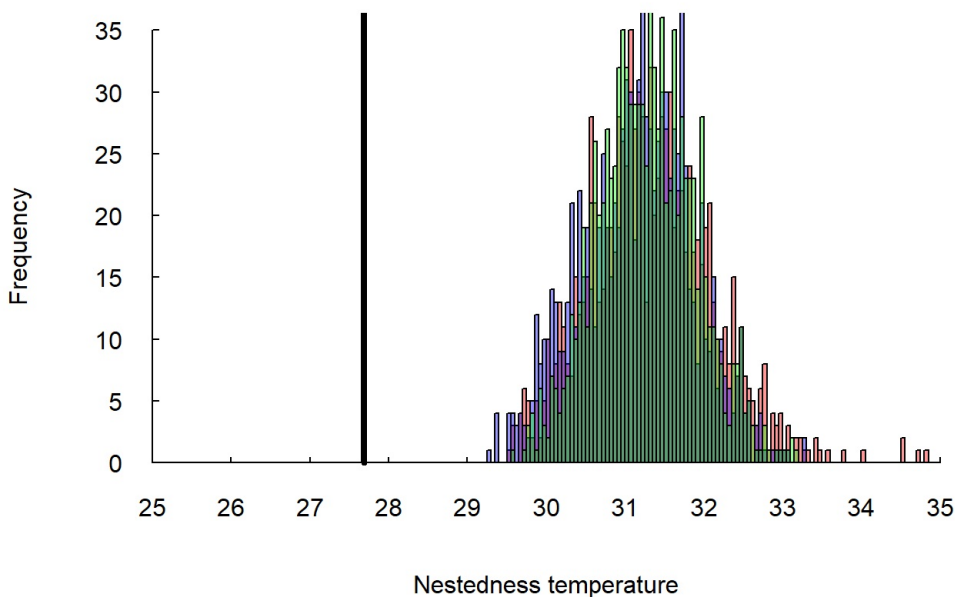
#Combine the three chains
rgbis<-c(rgbisA,rgbisB,rgbisC)

#Calculate the nestedness of networks in the reference distribution
null_nns<-lapply(rgbis,vegan::nestedtemp)

#Store the test statistic in a vector
null_nns_2<-numeric()
for(i in 1:length(null_nns)){
  null_nns_2[i]<-null_nns[[i]]$statistic
}

#Plot histogram (separate colour for each reference distribution and line for observed nestedness)
hist(null_nns_2[1:1000],col=adjustcolor("red",0.4),breaks=seq(25,35,0.05),main="",xlab="Nestedness temperature",xaxt="n",yaxt="n")
axis(side=2,pos=25,at=seq(0,35,5),las=1,tcl=0.2)
axis(side=1,pos=0,at=seq(25,35,1),las=1,tcl=0.2)
hist(null_nns_2[1001:2000],col=adjustcolor("blue",0.4),breaks=seq(25,35,0.05),add=TRUE,xaxt="n",yaxt="n")
hist(null_nns_2[2001:3000],col=adjustcolor("green",0.4),breaks=seq(25,35,0.05),add=TRUE,xaxt="n",yaxt="n")
lines(x=c(nn$statistic,nn$statistic),y=c(0,40),lwd=4)

```



By adding a line to the histogram to indicate the nestedness of their observed bipartite network, it is fairly apparent that small groups are more nested within larger groups than expected by chance. However, just to be careful they test this formally.

```
sum(c(null_nns_2,nn$statistic)<nn$statistic)
```

```
## [1] 0
```

The calculated P value is between 0.025 and 0.975 indicating that (assuming an alpha level of 0.05) there is no evidence that the nestedness of the observed network differs from that expected at random.

Team 3 are disease ecologists. They are concerned that the social system of this small population, in particular the fact it forms tight social groups, makes it vulnerable to emergent diseases. Because they therefore want to account for nondyadic (higher-order) interactions they compare insights from a **hypergraph** and dyadic network modelling approach.

An important note here is that the incidence matrix of a hypergraph is the same object as the bipartite network, so the difference is more conceptual than practical when we code the example below.

The reason that **team 3** want to utilize a hypergraph approach is that transmission dynamics can often be different in higher-order versus dyadic interactions. For more information on why this is the case we would refer readers to the papers “Capturing complex interactions in disease ecology with simplicial sets” (<https://doi.org/10.1111/ele.14079>) and “The physics of higher-order interactions in complex systems” (<https://www.nature.com/articles/s41567-021-01371-4>)

Here **team 3** assume the same non-linear dose-response curve as use in “Capturing complex interactions in disease ecology with simplicial sets” (<https://doi.org/10.1111/ele.14079>) and model disease dynamics accordingly. They simulate transmission dynamics through the hypergraph representation and then compare the results to the dyadic network representation.

First they define their transmission functions, first the hypergraph version and then the dyadic network version.

```
timestep_hyp<-function(classes,inc_mat,timer,dose){
  t_c<-classes

  for(i in 1:nrow(t_c)){
    if(t_c$S[i]==1){
      t_vec<-inc_mat[inc_mat[,i]==1,]
      t_vec[i]<-0
      classes$I[i]<-rbinom(1,1,1/(1+exp(-10*(sum(t_c$I*t_vec)*dose-0.5)))-1/(1+exp(-10*(-0.5))))
    }
  }

  classes$S[classes$I==1]<-0

  timer[is.na(timer)==FALSE&timer>0]<-timer[is.na(timer)==FALSE&timer>0]-1
  timer[classes$I==1&t_c$I==0]<-5

  classes$I[is.na(timer)==FALSE&timer==0]<-0
  classes$R[is.na(timer)==FALSE&timer==0]<-1

  return(list(classes,timer))
}

###

timestep_net<-function(classes,network,timer,dose){
  t_c<-classes

  for(i in 1:nrow(t_c)){
    if(t_c$S[i]==1){
      t_sum<-sum(t_c$I*network[i,])
      classes$I[i]<-rbinom(1,1,1-(1-(1/(1+exp(-10*(dose-0.5)))-1/(1+exp(-10*(-0.5))))))^t_sum
    }
  }

  classes$S[classes$I==1]<-0

  timer[is.na(timer)==FALSE&timer>0]<-timer[is.na(timer)==FALSE&timer>0]-1
  timer[classes$I==1&t_c$I==0]<-5

  classes$I[is.na(timer)==FALSE&timer==0]<-0
  classes$R[is.na(timer)==FALSE&timer==0]<-1

  return(list(classes,timer))
}
```

To better illustrate the example, we are going to assume that **team 3** have access to a longer-term data set (30 timesteps) than the other teams. They create a list with the network for each timestep.

```
groups<-interaction_generation_simul(id=id,adj_mat=adj_mat,mean_group_size=2,
                                   n_ts=30,float=0.000000001,pm=50,pow=4)

gbi_full<-groups[[1]]
day<-groups[[2]]

day_nets<-list()
for(i in 1:max(day)){
  day_nets[[i]]<-asnipe::get_network(gbi_full[day==i,])
}
```



```

infe_d<-numeric()

dose<-0.25

for(reps in 1:1000){

  classes<-data.frame(S=rep(1,50),I=rep(0,50),R=rep(0,50))
  classes[1:5,1]<-0
  classes[1:5,2]<-1

  timer<-c(rep(5,5),rep(NA,45))

  for(i in 1:max(day)){
    res<-timestep_net(classes=classes,network=day_nets[[i]],timer=timer,dose=dose)
    classes<-res[[1]]
    timer<-res[[2]]
  }

  infe_d[reps]<-sum(colSums(classes)[2:3])
}

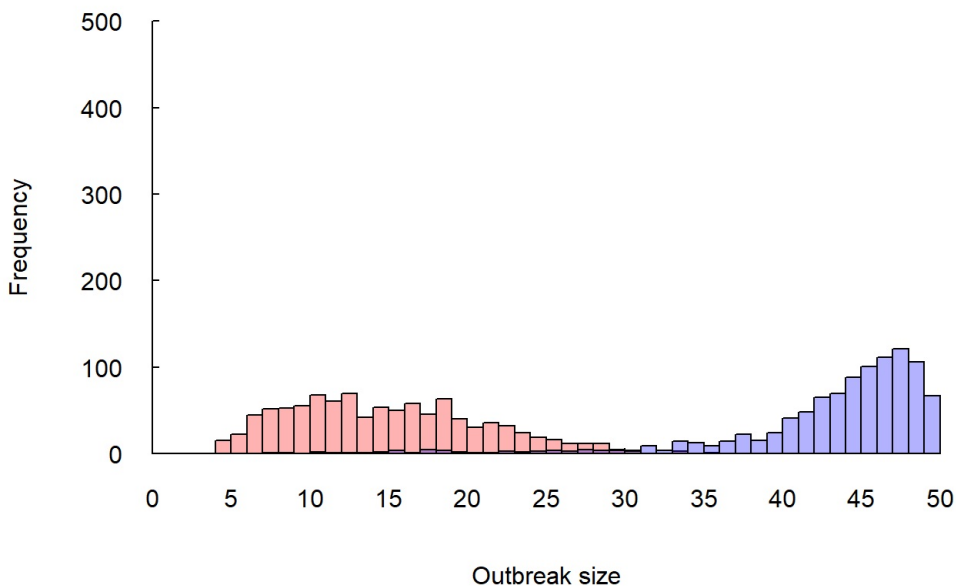
```

They then graphically compare the size of the outbreaks when accounting for nondyadic interactions (blue) with the the dyadic representation (red)

```

hist(infe_d,col=adjustcolor("red",0.3),breaks=seq(0,50,1),ylim=c(0,500),main="",xlab="Outbreak size",yaxt="n",xaxt="n")
axis(side=2,pos=0,at=c(0,100,200,300,400,500),las=1,tcl=0.2)
axis(side=1,pos=0,at=seq(0,50,5),las=1,tcl=0.2)
hist(infe_h,col=adjustcolor("blue",0.3),breaks=seq(0,50,1),add=TRUE,yaxt="n",xaxt="n")

```



And finally, they can also examine the probability the outbreak is larger when nondyadic interactions are accounted for using a resampling approach.

```

test_stat<-0
for(i in 1:10000){
  test_stat<-test_stat+(sample(infe_h,1)>sample(infe_d,1))
}
test_stat<-test_stat/10000
test_stat

```

```
## [1] 0.9856
```

We can see that in this case using the hypergraph explanation is important in forecasting the expected disease dynamics when this particular pathogen emerges in the population.