Network analysis of SDG interlinkages

Jonathan H.P. Dawes

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK

1 Introduction

This chapter aims to introduce and explain a number of general methods and ideas from network science, illustrating their use on two specific interlinkage matrices for the Sustainable Development Goals (SDGs). Formally a network comprises (i) a set of nodes (or vertices) and (ii) a set of edges between them. For the SDGs, the set of nodes could be chosen to be either the 17 goals, or the 169 separate targets, or subsets of these, as is commonly done when discussing a 'nexus' such as water–energy–food (Weitz, Nilsson & Davis, 2014; Yillia, 2016; Fader et al, 2018; van Noordwijk et al, 2018; Putra, Pradhan & Kropp, 2020). Edges describe interactions at the relevant level and carry varying kinds of information. Generally speaking, networks have (at least) three independent properties associated with each edge: (i) edges have direction: there is a direction of influence along each edge in the network; (ii) edges have weights: different interactions have different strengths; and (iii) edges have signs: edges can be negatively weighted (corresponding to trade-offs) or positively weighted (corresponding to co-benefits).

Depending on the data source and construction methodology, an SDG network may not exhibit all three features. Most obviously, networks based on correlations between data sources alone result in undirected networks since correlation does not in itself confer a direction of influence or causation. The fact that the construction and interpretation of a network may indeed rule out one of these properties is entirely natural; for example in disease transmission networks if the edges are associated with probabilities of transmission, these will all be non-negative quantities and so such edges do not have property (iii) above – they do not have signs.

From the text of the original resolution adopted by the UN General Assembly onwards, the Sustainable Development Goals (SDGs) have been framed as one "integrated and indivisible" agenda; see paragraphs 18 and 55 in particular of the resolution text (UN General Assembly, 2015). Yet the SDG agenda comprises 169 separate, and highly diverse, targets combined into 17 overarching goals. This tension between the unity of the whole and the diversity of the component parts has, from the launch of the SDGs onwards, made a description in terms of a network particularly appealing as a mode of thinking about Agenda 2030 (Le Blanc, 2015). A particular benefit is that it allows the application of quantitative methods of analysis of the SDGs, with the conceptual clarity that quantitative analysis offers.

With its focus on network science, this chapter naturally avoids discussing a number of important societal and political aspects of the SDGs. Two of these, for example, are (i) which sets of policy actors (government, business, third sector organisations, individual citizens) are enabled to carry out, or are responsible for, SDG implementation, and (ii) the relation between SDG interlinkages and policy coherence for sustainable development. The debate on the extent to which policy actions can be measured to have additional positive or negative additional influences (often referred to as 'co-benefits' and 'trade-offs', respectively, in the SDG literature) is clearly extremely important in understanding SDG interactions, but is omitted here.

1.1 Networks

Network-type problems arise in many fields of social and natural science and engineering, for example bioinformatics, behavioural ecology, national infrastructure, and telecommunications (Newman, 2018). Methods appropriate to each of these have been particularly well-studied over the last two decades; of course in many cases, notably in social sciences, the roots of these methodologies lie much deeper (Bonacich, 1972; Wasserman & Faust, 1994). More recent developments in network science, which we do not comment on further here but which clearly offer exciting directions for future research, include the analysis of multiplex (or multilayer) networks (Bianconi, 2018), in which the same collection of nodes may be connected by edges of different kinds (for example one layer corresponding to social interactions and a second layer corresponding to travel times between locations), and the analysis of networks in which interactions involve not just pairs of nodes, but 'higher-order' effects such as being reinforced in interactions between triples of nodes (Arrigo, Higham & Tudisco, 2020).

The richness and variety of network science tools means that there are challenges in selecting appropriate tools, with transparent methodologies, so that the results can be interpreted in the specific field. With this caveat, it is notable that systems thinking in general provides a flexible bridge between quantitative and qualitative insights into complex phenomena. Moreover, complex systems ideas of emergence and of phenomena that reach across scales indicate that it is necessary to apply this kind of thinking in order to to show how system level outcomes result from smaller-scale interactions.

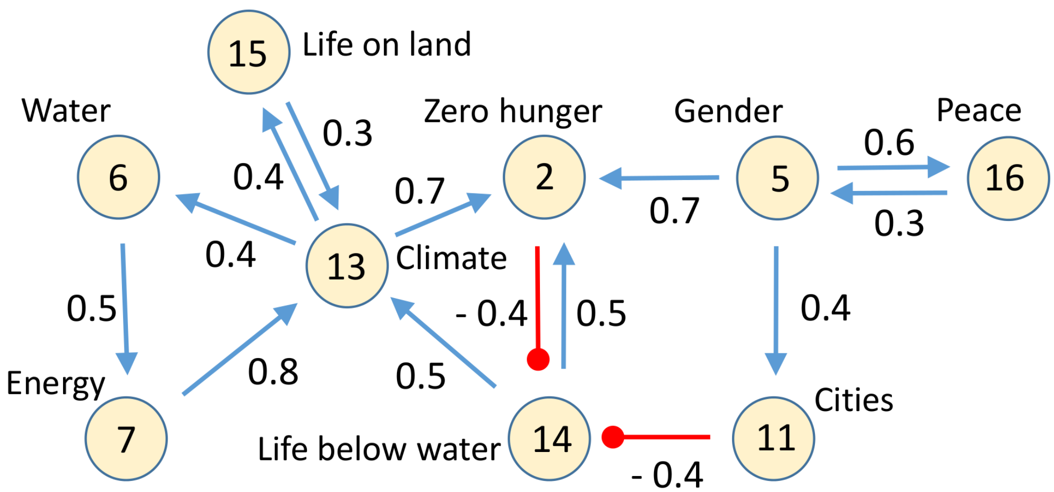


Figure 1: An example network for a subset of nine of the SDGs, with interlinkages indicated by the arrows. An arrow $j \to i$ indicates the positive influence of (progress on) SDG j on SDG i. A line with a round end indicates a negative influence of SDG j on SDG i. Adapted under CC-BY-4.0 from figure 2 in Dawes (2020).

Figure 1 shows an illustrative example of a network defined by a collection of nodes (here, a subset of nine SDGs), together with directed edges (arrows) indicating directions of influence between pairs of nodes. Arrowheads indicate positive influences, while round dots indicate negative influences. Each edge has a positive or negative weight associated with it; different edges have different weights. For convenience we assume that (by rescaling) the weights will all lie between $+1$ and -1. Several points of interest can be clearly seen from figure 1, such as the mutually reinforcing interlinkages between SDGs 6, 7 and 13 on the left-hand side of the figure, the negative influences of SDGs 2 and 11 on SDG 14, and more subtly, the fact that while SDGs 5 and 16 reinforce each other, neither receives any additional positive influence from the remainder of the network. But the overall net effect of the interlinkages is harder to quantify from figure 1; our need to understand the overall system-level response motivates the tools that we discuss in this chapter.

As a first step we encode the weighted directed influences into an *adjacency matrix* A which is simply a square array of the edge weights, with the array entry A_{ij} being the weight of the edge from node j to node i. Although this might seem to be slightly 'backwards' as a definition, it helps later on to ease the mathematical notation. The entry A_{ij} corresponds to the i^{th} row and the j^{th} column. Hence the adjacency matrix A corresponding to figure 1 is

$$\begin{array}{r} 2 \\ 5 \\ 6 \\ 7 \\ 11 \\ 13 \\ 14 \\ 15 \\ 16 \\ \\ \end{array} \begin{array}{c} \left(\begin{array}{rrrrrrrrr} 0 & 0.7 & 0 & 0 & 0 & 0.7 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.8 & 0 & 0 & 0.5 & 0.3 & 0 \\ -0.4 & 0 & 0 & 0 & -0.4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right), \\ \begin{array}{rrrrrrrrr} 2 & 5 & 6 & 7 & 11 & 13 & 14 & 15 & 16 \end{array} \end{array} \quad (1)$$

where we have included row and column labels to indicate precisely how the network nodes correspond to the rows and columns of the matrix. Note, for example the two negatively-weighted edges in the seventh row, corresponding to SDG 14, due to the negative influences on SDG 14 from SDGs 2 and 11. The adjacency matrix corresponds exactly to the 'node and edge' pictorial version of the network and one can be reconstructed immediately from the other. For larger networks with more edges we will not write out the numerical values but present a grey-scale 'heatmap' of the adjacency matrix, as in figure 2. This is an easier way to present a greater amount of information when the network diagrams become more complicated and the adjacency matrix has more non-zero entries. When we turn to consideration of adjacency matrices for 16 or all 17 SDGs, the adjacency matrix expands in size to become 16×16 or 17×17 respectively, but this is the only change in its interpretation.

1.2 Network data and quantitative approaches

Although not the main focus of this chapter, a note on the possible underlying sources of information for the construction of networks is appropriate here. SDG interlinkage networks are typically constructed from one of (broadly speaking) five possible sources of data: expert analysis; linguistic approaches (for example the identification of shared keywords across the literature); literature reviews and meta-analysis; statistical analyses, such as the analysis of correlations in time series for SDG indicators (see for example Pradhan et al, 2017); and the output of computational models, most obviously Integrated Assessment Models (van Soest et al, 2019) or bespoke models such as the iSDG framework (Pedercini et al, 2019). This five-fold classification and the need to integrate systemic features such as a gender perspective into this variety of analytical approaches are discussed further by Pollitzer (2022a,b). A slightly more detailed but essentially similar typology of data sources is given by Bennich, Weitz & Carlsen (2020) who propose seven types of data source and nine groups of analytical approaches, based on a sample of 70 papers from the academic literature. It is also of interest to note that some SDG interlinkage studies combine aspects of multiple approaches, for example the interlinkage matrices constructed by IGES (Zhou & Moinuddin, 2017, 2019; Dawes, Zhou & Moinuddin, 2022) that combine expert analysis which determines which interlinkages exist *a priori*, with data-driven correlation analysis at a country-specific level which then determines the strength of those interlinkages. A similar combination of data and expert opinion was used by Anderson et al (2021). Such interlinkage matrices can capture asymmetric interactions (via the expert opinion) while using correlation measures that are in themselves symmetric.

While data-driven approaches might appear to be the most preferable, given the concerns that will always exist around information gleaned through more sociological sources such as expert analysis, it needs to be borne in mind that data-driven methods have their own limitations and that these are often serious. The level of complexity that over 200 indicator time series bring, in their attempt to capture progress target-by-target is compounded by at least three significant effects. Firstly, the SDG indicator measure may only be a proxy for progress on a particular target, because the data required to monitor direct progress is not sufficiently widely available or sufficiently reliable. Secondly, the SDG target in question is often not framed sufficiently precisely that the relevant statistical indicator can be unambiguously defined and so may vary in its interpretation by different national statistical agencies. Thirdly, the data required may simply not be available and so the indicator time series cannot be constructed for the time frame or national coverage required. This is a particularly important issue in respect of SDG 5 where gender-disaggregated statistics are in embarrassingly short supply worldwide.

In addition to these three effects, there is a more subtle issue, which perhaps is a specific instance of the more general observation that all data sources have their limitations: data-driven approaches tend to be historical, and describe how progress arose in the past, but with shifts in policy or politics, these may well not be any guide to the future. In such a case, expert analysis for example may provide a more appropriate framework, and describe one of a range of possible futures, but help to decide priorities and actions that enable it. The interpretation of SDG interlinkage matrices naturally will depend on the data used in their construction.

Our focus here on quantitative, mathematical, methods demands that the connections between nodes must be numerically quantified. How one moves from qualitative to quantitative is often open to criticism and so must always be explicitly defined. Ideally the sensitivity of the results to changes in that part of the process should also be fully explored. One explicit relation between qualitative and quantitative views of SDG interlinkages that has gained popularity is the seven-point scale proposed by Nilsson, Griggs & Visbeck (2016) and presented in table 1. As discussed above, behind this scale sits, implicitly or explicitly, a sense of the scope of likely, or perhaps merely possible, policy actions and the effects that a policy action aimed at one SDG might have on others, either for better (a 'co-benefit') or worse (a 'trade-off').

Score	Name	Explanation
+3	Indivisible	Inextricably linked to the achievement of another goal
+2	Reinforcing	Aids the achievement of another goal
+1	Enabling	Creates conditions that further another goal
0	Consistent	No significant positive or negative interactions
-1	Constraining	Limits options on another goal
-2	Counteracting	Clashes with another goal
-3	Cancelling	Makes it impossible to reach another goal

Table 1: The seven-point scale proposed by Nilsson, Griggs & Visbeck (2016) for scoring the influence of one specific SDG or target on another.

It is of interest to note that this scale, while numerical, may not actually be amenable to numerical manipulation: for example if there are two separate effects that would lead to a linkage being scored as either +2 or -1, a net score of 0.5 (the arithmetic mean) or perhaps +1 as a compromise, might not be appropriate if the two underlying effects operated over different timescales or had different probabilities of arising.

1.3 Complete SDG interlinkage networks

Having used the relatively simple example network presented in figure 1 as an initial example, we illustrate the concepts developed in this chapter with two complete goal-level interlinkage matrices for the SDGs. The first of these was generated in Dawes (2020) from an expert analysis (ICSU-ISSC, 2015) and is the network from which the example presented in figure 1 was abstracted; in what follows we therefore use the complete version rather than this simplified verson of the network. The expert analysis was conducted by the International Science Council in 2015 when it was known as the International Council for Science (ICSU), in partnership with the International Social Science Council (ISSC); we therefore refer to this as the ICSU Report. Although based on a qualitative analysis of policy interlinkages rather than historic data, this is one of the few reports to treat all the SDGs identically (but not quite – it omits SDG 17). Dawes (2020) describes in detail the methodology through which the expert opinions were turned into a quantitative cross-impact matrix. Essentially this involved interpretation of the direction of interlinkages from the report's text, combined with an indication of the strength of an interlinkage given by the number of targets in the SDG that was being influenced that were mentioned as being impacted by progress on the influencing SDG. The initial example presented in figure 1 is in fact the subset of the ICSU network consisting of high-weight edges that form closed loops in the network; this can be seen by comparing the non-zero entries in the adjacency matrix (1) with the very light grey and very dark grey cells in figure 2(a).

The second interlinkage matrix we use is based on the Global Sustainable Development Report (GSDR) 2019 (GSDR, 2019; Pham-Truffert, Rueff and Messerli, 2019; Pham-Truffert et al, 2020). This report carried out a literature survey of 177 global scientific assessments, UN flagship report and scientific articles on SDG interlinkages, wherever possible looking at the level of targets; here we consider their results aggregated to goal-level. A hand-coding of statements in these articles resulted in a set of 4,976 separate positive and 782 negative interactions. Although there is considerable value in preserving the distinction between the positive and negative interactions separately, as discussed in more detail elsewhere (Dawes, 2022), for reasons of brevity we will discuss in this chapter only the 'net interlinkage' matrix obtained by summing positive and negative interlinkages.

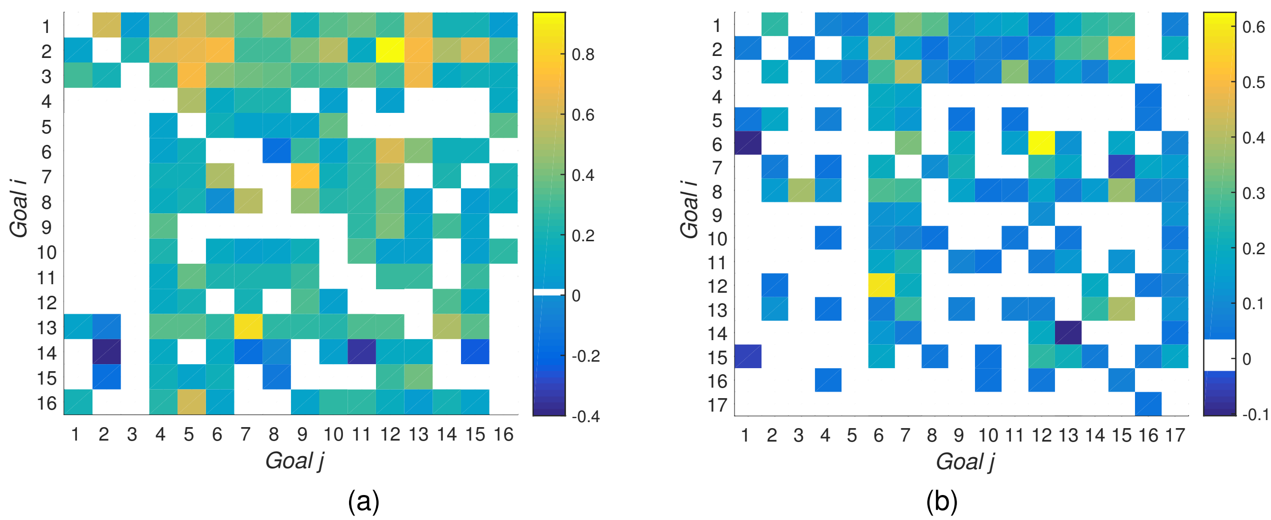


Figure 2: Heatmaps showing the distribution of positive and negative interlinkages in the SDG interaction matrices used as examples here. (a) The 16×16 ICSU matrix (SDG 17 was omitted from their analysis); (b) The 17×17 GSDR 2019 interlinkage matrix. Each cell corresponds to an interlinkage: the SDG labelling the column (j) influences the SDG labelling the row (i). White space indicates that the interlinkage is close to zero; color bars to the right give the scaled strength of interlinkages in each case. Adapted under CC-BY-4.0 from figure 1 in Dawes (2022).

Figure 2 shows the interlinkage matrices derived from the ICSU and GSDR reports. Interlinkages were scaled to lie in the range [-1, 1] and those close to zero are replaced by white space to improve the readability of the figures. Diagonal entries were removed from the GSDR matrix to make this more directly comparable with the ICSU matrix; removing these entries turns out to have only a very minor effect on the overall results (Dawes, 2022). In both cases we observe that almost every SDG influences SDGs 1 – 3, as shown by the densely populated top three rows of each plot. But there are fewer influences from SDGs 1 to 3 on the remaining goals, as is shown by the relatively sparsely coloured first three columns on the left hand side of each plot. This highlights what appears to be a fundamental asymmetry in Agenda 2030: the first three SDGs are those on which most attention is focused, with SDGs 4 to 17 in some sense playing a subordinate role, supporting the fundamental development agenda, and continuation from the Millennium Development Goals, that SDGs 1 – 3 represent.

The interlinkage matrices in figure 2 clearly differ from each other, and this highlights the obvious fact that interlinkage matrices constructed from different data sources vary, and often vary considerably. The point is not to find a single interlinkage matrix that is somehow the most appropriate, or optimal, one, but to be able to link aspects of the detailed construction of any one interlinkage matrix with the system-level implications of those goal-by-goal, or indeed target-by-target influences. Indeed, even being forced to address issues such as the directionality of interlinkages, often helps to clarify what a particular interlinkage matrix is attempting to capture.

1.4 Chapter organisation

The remainder of this chapter is organised as follows. In section 2 we explain and apply the key idea of *centrality*, that is, the relative importance of different nodes in a network. Of the many different ideas of centrality we focus on the idea known as 'eigencentrality' (Bonacich, 1972, 2007; Newman, 2018; Bali Swain & Ranganathan, 2021) which in a strict sense does not apply to complex networks with both positive and negative interlinkages, but has an alternative and extremely useful interpretation making it an important and applicable system level statistic (Dawes, 2020).

Already above we have noted that the robustness of any network-based measures is an important aspect of these quantitative approaches. In section 3 we discuss two possible measures of the sensitivity of the eigencentrality measure to perturbations to the network (that is, the addition or removal of new interlinkages). This also leads naturally to methods to answer the question of where best to introduce new interlinkages in order to improve the self-reinforcements already present in a network.

Section 4 discusses the third idea: the notion of hierarchy among an interlinkage network for the SDGs. Measures of overall hierarchy and directionality in complex networks have their roots in the computation and analysis of ecosystems, for example food webs, but lend themselves also to the systemic analysis and quantification of prioritisation between the SDGs. Prioritisation of SDGs that lie further 'upstream' of others should enable those goals to be met while at the same time allowing benefits to flow through the network and so allow the whole system to benefit. In contrast, prioritisation of SDGs that lie far 'downstream', and hence have much lower levels of influence on other goals, would not allow all SDGs to be met. We develop and discuss measures of hierarchy that help to identify points of leverage and maximum 'downstream benefit' to other SDGs in the network. We conclude in Section 5 and offer perspectives and directions for further research.

2 Centrality measures

A common observation in network science is that some nodes appear to be more important, or more 'central' to the network than others. The most fundamental notion of centrality is simply to count the edge weights of the edges by which other nodes are connected to the node under consideration. This gives rise to the centrality measures

$$k_i^{\text{in}} = \sum_{j=1}^{n} A_{ij} \qquad \text{and} \qquad k_i^{\text{out}} = \sum_{j=1}^{n} A_{ji} \tag{2}$$

which are the (weighted) in-degree k_i^{in} and out-degree k_i^{out} of node i, respectively. We note that in situations where edges weights may be negative as well as positive, it might be more appropriate to take the absolute value of the edge weights A_{ij} in order to avoid cancellation between positive and negative contributions. Compact expressions for the vectors of in-degrees and out-degrees can be written making use of the vector $\mathbb{1}$ which denotes the vector $(1, 1, \ldots, 1)$ in which every element is one. Then $\boldsymbol{k}^{\text{in}} := A\mathbb{1}$ and $\boldsymbol{k}^{\text{out}} := A^T\mathbb{1}$ where A^T denotes the transpose of the matrix A, that is, the matrix in which the $(i, j)^{\text{th}}$ element is the entry A_{ji}.

Pham-Truffert et al (2020) refer to nodes that have a large in-degree as 'buffers' since in some sense they serve to combine the effects of many different nodes together, and they refer to nodes that have a large out-degree as 'multipliers' as they may propagate the influence of a node to many other parts of the network. The total degree $k_i = k_i^{\text{in}} + k_i^{\text{out}}$ of node i is a natural measure of the relative importance of node i in the network; this is commonly referred to as the *degree centrality* of the node.

Other popular and long-studied measures of centrality, which we will not consider further here, include 'betweenness centrality' (Freeman, 1977) and 'closeness centrality' (Bavelas, 1950). To compute betweenness centrality we first compute the set of shortest paths between every pair of nodes in the network. The betweenness centrality of a node is then proportional to the number of shortest paths that pass through that node. Closeness centrality is the average length of the shortest path between the given node and all other nodes in the network. Both definitions are most straightforward to apply to unweighted, undirected and connected networks; various generalisations and improvements account for cases in which the network is either weighted, directed, or disconnected (a network is disconnected if its set of nodes divides into two subsets for which there are no edges linking any pair of nodes chosen so that one node lies in each subset).

2.1 Eigencentrality

While simple, the concept of degree centrality has an inherent drawback – it is a purely local calculation. By 'local' we mean that degree centrality counts the numbers of direct neighbours of a node but does not make any allowance for how connected those nodes themselves might be. A node with a large number of 'unimportant' neighbours is perhaps less important itself than a node with a smaller number of 'very important' neighbours. A more robust measure of importance could therefore be obtained differently, through a slightly self-referential definition: the importance of a node is given by the weighted sum of the importance of the nodes to which it is connected. Mathematically we can define the importance v_i of node i to be given by the weighted sum

$$v_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{ij} v_j \tag{3}$$

where λ is a 'normalisation' or rescaling parameter. This definition appears to be unfortunately circular since it demands knowing the importances v_j of the nodes j to which node i is connected. One can imagine some kind of iterative scheme, starting with estimated values for the v_i and then re-calculating them according to the formula (3) until they converge. Such an approach indeed works (for almost all initial choices of the v_i) although it involves estimating the value of λ as well. In more mathematical language, λ is the largest, or 'leading' eigenvalue of the matrix A and $\mathbf{v} = (v_1, \ldots, v_n)$ is the leading eigenvector. The calculation of λ and $\mathbf{v}$ is numerically straightforward in many software packages. Further mathematical remarks are given in Box 1.

Box 1. Eigenvalues and eigenvectors. In the generic case, an $n \times n$ matrix A has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$ each with a corresponding eigenvector $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}$. Each eigenvector is a column vector with n elements. We order the eigenvalues so that $Re(\lambda_1) \geq Re(\lambda_2) \geq \cdots \geq Re(\lambda_n)$ where Re denotes the real part of the eigenvalue λ_i which is possibly a complex number even when the matrix A is real.

The eigenvalues and eigenvectors are defined by the equation $A\mathbf{v}_i = \lambda_i \mathbf{v}^{(i)}$. In practice we usually find eigenvalues first, by solving the equation $P(\lambda) := \det(A - \lambda I) = 0$, where $\det$ denotes the determinant of the matrix, and I denotes the $n \times n$ identity matrix. $P(\lambda)$ is a polynomial of degree n in λ and so (generically) has n distinct roots which are the eigenvalues $\lambda_1, \dots, \lambda_n$. For each eigenvalue one can then find the column vector $\mathbf{v}$ that satisfies $(A - \lambda_i I)\mathbf{v}^{(i)} = \mathbf{0}$. Eigenvectors are typically scaled ('normalised') so that the sum of the squares of their entries is one; if this is carried out then each entry lies in the range from -1 to $+1$.

For a typical matrix A that has no negative entries the Perron–Frobenius Theorem (Meyer, 2000, Chapter 8) guarantees that the largest eigenvalue λ_1 is real and positive and it has a corresponding eigenvector $\mathbf{v}^{(1)}$ that has no negative entries. In this case, for typical positive initial guesses, iterations of equation (3) when λ is set to the value λ_1 converge to $\mathbf{v}^{(1)}$.

The elements of the leading eigenvector $\mathbf{v}^{(1)}$ define a centrality measure describing the relative importance of each of the nodes; this is known as eigenvector centrality, abbreviated sometimes to *eigencentrality*. Eigencentrality is a remarkably old concept, dating back at least as far as the article by Landau (1895) on scoring in chess tournaments; further historic references include Leontief (1941) and Seeley (1949). If the matrix A has no negative entries then the eigencentrality scores for each node lie in the range zero to one, assuming the normalisation described in Box 1 is carried out.

2.2 Interpretation

Eigencentrality as an importance measure is related to the in-degrees of the nodes rather than their out-degrees in the sense that, for a directed network (as we consider in this chapter) a large eigencentrality score relates to a high accumulation of co-benefits provided by other nodes. Therefore, in the context of SDG interlinkage networks, eigencentrality according to the definition (3) measures the extent to which a node receives co-benefits (or trade-offs) from other nodes, rather than the extent to which it generates these co-benefits and confers them onto other nodes. Therefore, the elements of $\mathbf{v}^{(1)}$ provide a measure of the extent to which nodes are 'sinks', rather than 'sources'.

While the Perron–Frobenius Theorem referred to in the last paragraph within Box 1 is mathematically well-defined only for non-negative matrices, it is often the case that a network with only a relatively small number of negative links will also have a leading eigenvector that has all entries non-negative; in such a case it is tempting to continue to interpret the leading eigenvector as a centrality measure. But there is an additional context for interpreting the leading eigenvector – it is the dominant 'response' of the network when we consider the interlinkage network as dynamically generating reinforcements, or trade-offs between progress on different nodes over time. For example, taking the simplest possible case, consider the evolution equation

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \sum_{j=1}^{n} A_{ij} x_j \,, \tag{4}$$

which states that the rate of change in time t of a state variable x_i related to node i depends linearly on each of the state variables x_j at the nodes j that have influence on, and feed into, node i, mediated in each case by the strength and sign of the interaction between the two nodes as captured by the interlinkage A_{ij}. The solution $\boldsymbol{x}(t)$ to the differential equation (4) is dominated, apart perhaps over a short initial transient phase, by the form of the leading eigenvector because the leading eigenvector describes the mode of maximum growth rate – the eigenvalues $\lambda_1, \ldots, \lambda_n$ are the growth rates of these different modes of response, but $\mathbf{v}_1$ dominates since (by definition) λ_1 is larger than all the others. For completeness, we note that in special cases there might be other eigenvalues equal to λ_1 but this is not a generic behaviour and so in the interests of simplicity, and because it does not arise in practical examples of SDG interlinkage networks, we gloss over this point in our exposition here. Indeed, the examples below we will see that the leading eigenvalue is substantially larger than the remaining $n-1$ eigenvalues and so provides a good guide to the overall system dynamics.

To summarise, the leading eigenvector can be interpreted as the intrinsic mode of self-reinforcing growth of the state variables $\boldsymbol{x}(t) = (x_1(t), \ldots, x_n(t))$ caused solely by their interactions. In the context of this work, and conscious of many caveats around the simplistic nature of equation (4) and the coarse-grained representation this implies, we can view $x_1(t), \ldots, x_n(t)$ as the relative levels of progress made on each of the SDGs, interpreting the interlinkages as reinforcing, or constraining effects due to policy actions.

2.3 A measure of influence

We now briefly contrast the interpretation of eigencentrality above (in which it reveals which nodes benefit the most from positive interlinkages in the network) with a closely-related but distinct question: which nodes have the greatest influence on others, across the network? That is, instead of asking which nodes receive the greatest number of co-benefits (as 'sinks'), we ask which nodes provide the most co-benefits, as 'sources'. Let y_i denote the influence across the network of node i. As in the earlier discussion, we can frame an answer to deducing the relative levels of influence through a similar but subtly different self-referential definition: the influence of node i is given by the weighted sum of the influences of the nodes that node i is itself able to influence it. Mathematically we can define the influence y_i of node i to be given by

$$y_i = \frac{1}{\lambda} \sum_{j=1}^{n} y_j A_{ji} \,. \tag{5}$$

The key difference between (5) and (3) is the directionality of the interlinkages: note the term A_{ji} in (5) rather than A_{ij}. A node i that has no outgoing edges (i.e. where $k_i^{\text{out}} = 0$) will have zero influence on the network: this agrees with (5) since $k_i^{\text{out}} = 0 \implies y_i = 0$. However, a node with no outgoing edges may still have a positive in-degree $k_i^{\text{in}} > 0$ and hence $v_i > 0$. The distinction between the levels of 'influence' and 'benefits received' for each node is crucial for directed networks of the kind that we consider here.

Mathematically, (5) implies that $\mathbf{y}$ satisfies the equation $\mathbf{y}^T A = \lambda \mathbf{y}^T$. The row vector $\mathbf{y}^T$ is the left eigenvector corresponding to the eigenvalue λ, in contrast to $\mathbf{v}$ which is the (usual) right eigenvector. Equivalently, $\mathbf{y}^T$ is the right eigenvector of the transpose matrix A^T defined by swapping the elements of A so that $(A^T)_{ij}$ is the element A_{ji}.

2.4 Katz centrality

Finally, we briefly mention a centrality measure that 'interpolates' between eigencentrality and degree centrality (Katz, 1953). Consider the importance x_i of node i to be given by the weighted sum of the importances of its neighbours as in (3), with a weighting parameter α, but also including a constant term so that every node has importance at least one:

$$x_i = \frac{1}{\alpha} \sum_{j=1}^{n} A_{ij} x_j + 1 \,. \tag{6}$$

Written in matrix-vector terms this is equivalent to writing $\boldsymbol{x} = \alpha A \boldsymbol{x} + \mathbb{1}$ where $\mathbb{1}$ is the column vector having a one in every element. Re-arranging we obtain

$$\boldsymbol{x} = (I - \alpha A)^{-1} \mathbb{1} = (I + \alpha A + \alpha^2 A^2 + \alpha^3 A^3 + \cdots) \mathbb{1} \,, \tag{7}$$

where I denotes the $n \times n$ identity matrix having ones on the main diagonal and zeros in all other entries, and the power series expansion indicated by the ellipsis is valid for $0 \le \alpha < 1/\lambda_1$. The parameter α measures the 'discount rate' at which the contribution to the importance of node i from ever more distant nodes drops off as the distance through the network increases. For very small α only the immediate neighbours of node i really count, and so the centrality measure $x_i \approx 1 + \alpha k_i^{\text{in}}$ and we recover (apart from the constant 1 and the scaling by α which do not affect the relative values of the centralities) the in-degree centrality. In the opposite limit, as $\alpha \to 1/\lambda_1$ we recover the eigencentrality measure v_i.

Similarly, if we replace A by its transpose in (6) and (7) we obtain measures of the influence u_i of node i, with successive terms in the sum on the right-hand side of (7) corresponding to contributions from nodes at successive distances from node i. This is precisely what the analysis by Weitz et al (2018) aimed to capture in the analysis of a cross-impact matrix for a set of 34 SDG targets in a country-level analysis for the case of Sweden. Their overarching aim was to understand how to answer the question (as they posed it) *" If progress is made on target x, how does this influence progress on target y? "*. Weitz et al proposed the idea that the total influence of a node should be a combination of the direct ('first order' in their terminology) and indirect ('second order') influences. This combination of direct and indirect influences led Weitz et al to propose the measure of total influence $\mathcal{I}_i^{\text{total}}$ of a node i as a weighted sum of its out-degree k_i^{out} and the out-degrees of its neighbours, multiplied by the interlinkage strength A_{ij} (noting that they used the convention that A_{ij} is the strength of the edge $i \to j$ and so their A is our A^T):

$$\mathcal{I}_i^{\text{total}} = k_i^{\text{out}} + \frac{1}{2}\sum_{j=1}^{n} A_{ij} k_j^{\text{out}} . \tag{8}$$

In light of the discussion above we can see this as a truncation of the series expansion in (7), ignoring the $\alpha^2 A^3$ and higher-order terms, and with A replaced by A^T in order to compute influences rather than importances. In our notation, letting $\mathcal{I}_i^{\text{total}} = (u_i - 1)/\alpha$, and setting $\alpha = 1/2$, their definition in equation (8) above becomes

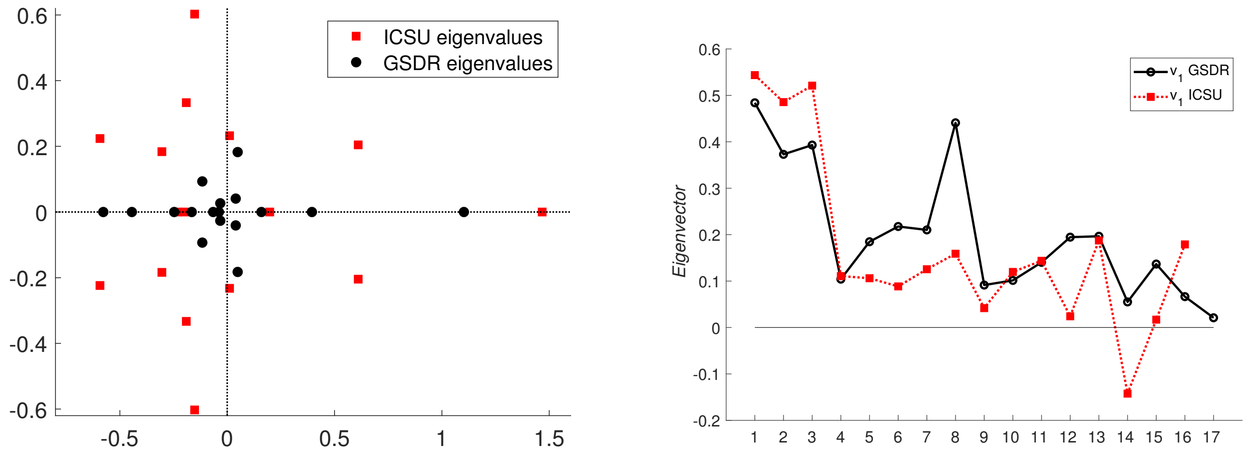
$$\begin{aligned} (\boldsymbol{u} - \mathbb{1})/\alpha &= A^T\mathbb{1} + \frac{1}{2}(A^T)^2\mathbb{1} = \left[A^T + \alpha(A^T)^2\right]\mathbb{1} , \\ \Longrightarrow \boldsymbol{u} &= \left(I + \alpha A^T + \alpha^2 (A^T)^2\right)\mathbb{1} , \end{aligned}$$

(using the fact that $I\mathbb{1} = \mathbb{1}$) which is precisely the truncation of (7) ignoring the A^3 term and all higher powers. Again, in the limit as $\alpha \to 1/\lambda_1$ (and including the higher-order terms in α) the influence measure $\boldsymbol{u}$ converges to the eigenvector-based influence measure $\mathbf{y}$ discussed previously.

2.5 Illustrations of centrality

It is numerically straightforward to compute the eigenvalues and eigenvectors for the interlinkage matrices shown in Figure 2; the results are given in Figure 3. Figure 3(a) plots the eigenvalues. In both cases there is a leading eigenvalue that appears far to the right of the remaining ones, leaving a significant gap between the leading eigenvalue and the next-largest, for both the ICSU eigenvalues (squares) and GSDR eigenvalues (round dots). This indicates that the behaviour of the networks is dominated by the form of their leading eigenvectors and that this fact is robust to perturbations of individual interlinkages. The form of the leading eigenvectors themselves is shown in Figure 3(b), again with the GSDR eigenvector in black and the ICSU one in red. The lines joining the dots are added just as a guide to the eye; the data points show the components of the eigenvectors for each of the SDGs. The ICSU eigenvector has only 16 components since the ICSU analysis omitted SDG 17.

The asymmetry noticed above is reflected in the high values of the eigenvector components for SDGs 1 – 3: these three SDGs are positively supported by many of the others which therefore act to promote greater progress on goals 1 - 3, in the dynamical sense, and for them to be identified as among the most important in a centrality sense. The GSDR matrix has a particularly high component also for SDG 8 (Decent work and economic growth) which is perhaps expected due to the relatively full row of interactions supporting SDG 8 in



(a) (b)

Figure 3: Eigenvalues and (right) eigenvectors for the ICSU and GSDR interlinkage matrices. (a) Eigenvalues of the adjacency matrices A in the two cases, shown in the complex plane (i.e. the horizontal axis corresponds to the real part of the eigenvalue and the vertical axis to the imaginary part). The leading eigenvalues are real and are located at approximately $\lambda_1 = 1.47$ for the ICSU matrix and $\lambda_1 = 1.10$ for the GSDR matrix (both to 2 decimal places). In both cases there are significant horizontal gaps between this leading eigenvalue and those that are next-largest. (b) The leading eigenvectors $\mathbf{v}^{(1)}$ for the ICSU and GSDR adjacency matrices, with the components for each SDG plotted against the SDG number on the horizontal axis. The vertical scale is arbitrary since eigenvectors are defined only up to a scale factor and are normalised here as described in Box 1 (so that the sums of the squares of the components add up to one). Adapted under CC-BY-4.0 from figure 2 in Dawes (2022).

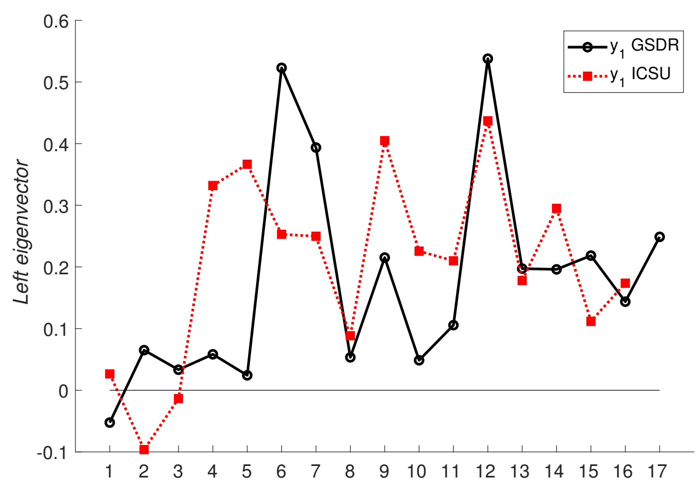


Figure 4: The components of the left eigenvectors $\mathbf{y}^{(1)}$ corresponding to the leading eigenvalue λ_1 for the ICSU and GSDR adjacency matrices, with the components corresponding to each SDG plotted against the SDG number on the horizontal axis. The vertical scale is arbitrary since eigenvectors are defined only up to a scale factor. Here, as is customary and convenient, we normalize so that the sum of the squares of the components adds up to one.

the interaction matrix in Figure 2(b). The corresponding row for the ICSU matrix in Figure 2(a) also strongly supports SDG 8 but in this case the support is stronger for other SDGs and so SDG 8 does not emerge as relatively better off. Most concerning is the negative component of the leading eigenvector for SDG 14 (Life below water) for the ICSU matrix. This suggests that the internal dynamics of the ICSU network would result in negative progress on SDG 14 when positive progress is made elsewhere. In terms of the interlinkage matrix there are two strongly negative links, from SDG 2, and from SDG 11, to SDG 14 – these are coloured dark blue in Figure 2(a) and point to significant trade-offs between the goals on zero hunger and sustainable cities, and progress on life below water.

Figure 4 reveals that the most influential sources of co-benefits are SDGs 12, 9, 5 and 4 (in that order) for the ICSU interlinkage network (and these in effect generate the co-benefits that result in greatest progress on SDGs 1-3 as revealed by figure 3b). For the GSDR network SDGs 12, 6 and 7 are the most influential in generating co-benefits for other SDGs (and from figure 3(b) we see that these benefits accrue most to SDGs 1-3 and 8 for the GSDR network). We see also that for the ICSU matrix SDGs 2 and 3 are a source of trade-offs within the network, while SDG 1 is the only SDG that is a source of (net) trade-offs for the GSDR interlinkage matrix.

3 Sensitivity to individual linkages

Given the importance of eigenvalues and eigenvectors in summarising the overall dynamics of an interlinkage network, it is important and useful to be able to relate them directly to variations in the microscopic individual links in the network. This enables us to explore, for example, whether and how we might be able to influence the overall dynamics through changes to one or a small number of linkages. This discussion of the sensitivity of centrality results to the addition of new network linkages is linked to long-standing and well-known perturbation results in applied mathematics and yields useful analytical results. It also yields readily to interpretations, building on the results of the previous section and the concepts of right and left eigenvectors introduced there. The discussion in this section is based on Dawes (2022).

First, we distinguish between two possible senses in which we might want to improve the co-benefits that the interlinkage matrix results in. Our first measure, which we term 'growth rate sensitivity' describes the change in the leading eigenvalue which is the overall rate at which interlinkages generate overall changes (which we hope are largely positive) in progress on the SDGs. The second measure, which we term 'equality sensitivity' describes the extent to which a change in a particular interlinkage affects the distribution across the SDGs of these self-reinforcing benefits so that progress becomes more equal across the whole set of SDGs.

3.1 Growth rate sensitivity

A suitable mathematical measure of the sensitivity of the growth rate to perturbations is to consider the rate of change of the leading eigenvalue λ_1 of the interlinkage matrix to perturbations in one element A_{ij} of the adjacency matrix. We therefore define the 'multiplier effect' sensitivity matrix S^{m} to be the $n \times n$ matrix whose (i, j) entry is $S^{\mathrm{m}}_{ij} := (1/\lambda_1)\partial\lambda_1/\partial A_{ij}$. The factor of $1/\lambda_1$ is included so that S^{m} computes relative changes in λ_1 rather than absolute changes. In the cases of interest here, S^{m}_{ij} turns out to have a simple expression in terms of the left and right eigenvectors ($\mathbf{y}^{(1)T}$ and $\mathbf{v}^{(1)}$, respectively) for the leading eigenvalue λ_1. A straightforward calculation, presented in Appendix B of Dawes (2022) and also summarised in Box 2, leads to the formula

$$S^{\mathrm{m}}_{ij} = \frac{1}{\lambda_1} \frac{y_i^{(1)} v_j^{(1)}}{\mathbf{y}^{(1)T}\mathbf{v}^{(1)}} \tag{9}$$

Thus S^{m}_{ij} is a measure of the sensitivity of the leading eigenvalue of the adjacency matrix to an increase in the interlinkage A_{ij}.

The growth rate sensitivity formula (9) has an intuitive explanation that builds on the interpretations of $\mathbf{y}^{(1)}$ as the sources of co-benefits (or trade-offs) in the network, and $\mathbf{v}^{(1)}$ as the resulting 'sinks' of co-benefits or trade-offs. Consider the two cases in figure 5 where we show, as thought-experiment, two nodes labelled i and j and part of their interlinkage network.

For simplicity, suppose that initially the nodes are not linked, and consider the effect on the network overall of introducing a new edge from j to i. In case (a), corresponding to figure 5(a), suppose that node i has only incoming edges and no outgoing ones: $k_i^{\mathrm{out}} = 0$. (We make no assumptions about the nodes that j is connected to; suppose j has both incoming and outgoing edges.) Then node i is unable to propagate co-benefits or trade-offs further through the network; column i in the interlinkage matrix must consist entirely of zeros: $A_{pi} = 0$ for all $p = 1, \ldots, n$. Then from (5) we conclude that we must have $y_i = 0$ for any left eigenvector $\mathbf{y}$. So in particular this is true for the leading left eigenvector $\mathbf{y}^{(1)}$ and hence $S_{ij}^{\mathrm{m}} = 0$ from (9) and so the addition of the new link $j \to i$ does not change the growth rate λ_1. Intuitively, if node i only receives benefits from other nodes and is not a source itself, then adding a new edge pointing towards node i does not change that situation, and so the network as a whole does not gain or lose.

In case (b), corresponding to 5(b), consider adding the new edge $j \to i$ when j has no incoming edges, only outgoing ones (and we suppose that i might have both incoming and outgoing edges). Then node j is unable to receive the results of co-benefits or trade-offs from other nodes: row j of the adjacency matrix contains only zeros: $A_{jq} = 0$ for all $q = 1, \ldots, n$. Hence, from (3) we see that v_j must be zero, that is, the j^{th} entry of any (right) eigenvector $\mathbf{v}$ must be zero. Then from (9) we see that $S_{ij}^{\mathrm{m}} = 0$ and so the addition of the new edge $j \to i$ does not change the growth rate for the network at all.

To summarise, the formula 9 can be justified intuitively as stating that, for a new edge $j \to i$ to affect the network dynamics, it is necessary that node i already propagates co-benefits or trade-offs through to other nodes in the network (that is, $k_i^{\mathrm{out}} > 0$, and that node j already receives co-benefits or trade-offs from other nodes (that is, $k_j^{\mathrm{in}} > 0$). The sensitivity of the network to the addition of the new edge $j \to i$ is proportional to the product of the level of co-benefits currently offered by node i and received by node j.

Figure 5: Illustrations of the two cases in which we would expect the growth rate sensitivity to the addition of an edge from node j to node i to be zero. (a) Suppose that j has a mixture of incoming and outgoing edges (not shown), but that i has only incoming edges, and that we add a new edge $j \to i$. (b) Suppose in this case that i has a mixture of incoming and outgoing edges (not shown), but that j has only outgoing edges, and that we add a new edge $j \to i$.

Box 2. Derivation of the sensitivity matrix S^{m}. Following Greenbaum et al (2020) here we summarise the derivation of (9) that describes the effect on the growth rate λ_1 of a perturbation to the interlinkage from j to i, that is, the entry A_{ij}. Let the matrix A have distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, with corresponding right and left eigenvectors $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(n)}$ and $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(n)}$ normalised so that $\|\mathbf{v}^{(i)}\| = \|\mathbf{y}^{(i)}\| = 1$ for all $i = 1, \ldots, n$.

Now consider the adjacency matrix A to be a function of a parameter ε that changes the interlinkage from j to i (this will be made explicit in a moment), writing $A(\varepsilon)$ where $A \equiv A(0)$. Suppose that $A(\varepsilon)$ has an eigenvalue $\lambda_1(\varepsilon)$ and a right eigenvector $\mathbf{v}^{(1)}(\varepsilon)$ that vary smoothly with ε, so that $A(\varepsilon)\mathbf{v}^{(1)}(\varepsilon) = \lambda_1(\varepsilon)\mathbf{v}^{(1)}(\varepsilon)$. Differentiating with respect to ε and setting $\varepsilon = 0$ we obtain

$$A'(0)\mathbf{v}^{(1)}(0) + A(0)\mathbf{v}^{(1)\prime}(0) \quad = \quad \lambda_1'(0)\mathbf{v}^{(1)}(0) + \lambda_1(0)\mathbf{v}^{(1)\prime}(0),$$

where $'$ denotes a derivative with respect to ε. Now multiply on the left by the left eigenvector $\mathbf{y}^{(1)T}(0)$ and observe that the second term on the left hand side will cancel with the second term on the right hand side since they are both $\lambda_1(0)\mathbf{y}^{(1)T}(0)\mathbf{v}^{(1)\prime}(0)$. Hence we obtain

$$\lambda_1'(0) \quad = \quad \frac{\mathbf{y}^{(1)T}(0)A'(0)\mathbf{v}^{(1)}(0)}{\mathbf{y}^{(1)T}(0)\mathbf{v}^{(1)}(0)}. \tag{10}$$

This expression holds for any perturbation to the interlinkage matrix A; we now specialise to the case that the perturbation is to a single network edge $j \to i$, for which the derivative matrix $A'(0)$ is just the outer product $A'(0) = \mathbf{e}_i\mathbf{e}_j^T$ where $\mathbf{e}_i$ is the column vector whose i^{th} entry is 1 and all other entries are zero. For this case, omitting the parameter dependence in the notation, the expression (10) becomes

$$\lambda_1' \quad = \quad \frac{y_i^{(1)}\, v_j^{(1)}}{\mathbf{y}^{(1)T}\mathbf{v}^{(1)}}$$

which describes the rate of change of the eigenvalue λ_1 with respect to the value of the network edge from $j \to i$. This result is well-known (Jacobi, 1846) and was re-derived more recently in the ecological literature (see equation (23) in Neubert & Caswell, 1997) in the context of ecosystem resilience.

Finally, the matrix S^{m} introduced in (9) is precisely the $n \times n$ matrix of rates of change with respect to perturbations in each element of A in turn, scaled by a factor of λ_1. We may define:

$$S_{ij}^{\mathrm{m}} \quad := \quad \frac{1}{\lambda_1}\frac{\partial \lambda_1}{\partial A_{ij}} = \frac{1}{\lambda_1}\frac{y_i^{(1)}\, v_j^{(1)}}{\mathbf{y}^{(1)T}\mathbf{v}^{(1)}}\,. \tag{11}$$

The factor of $1/\lambda_1$ is included in our definition so that S^{m} computes relative changes in the magnitude of the leading eigenvalue: this allows better comparison of the values of S^{m} computed from different matrices with different absolute values of λ_1.

3.2 Equality sensitivity

The second sensitivity measure is also easy to define in terms of eigenvectors, although some of the mathematical details are a little more complicated and are omitted here. We define the equality sensitivity S_{ij}^{eq} of the leading eigenvector $\mathbf{v}^{(1)}$ to be

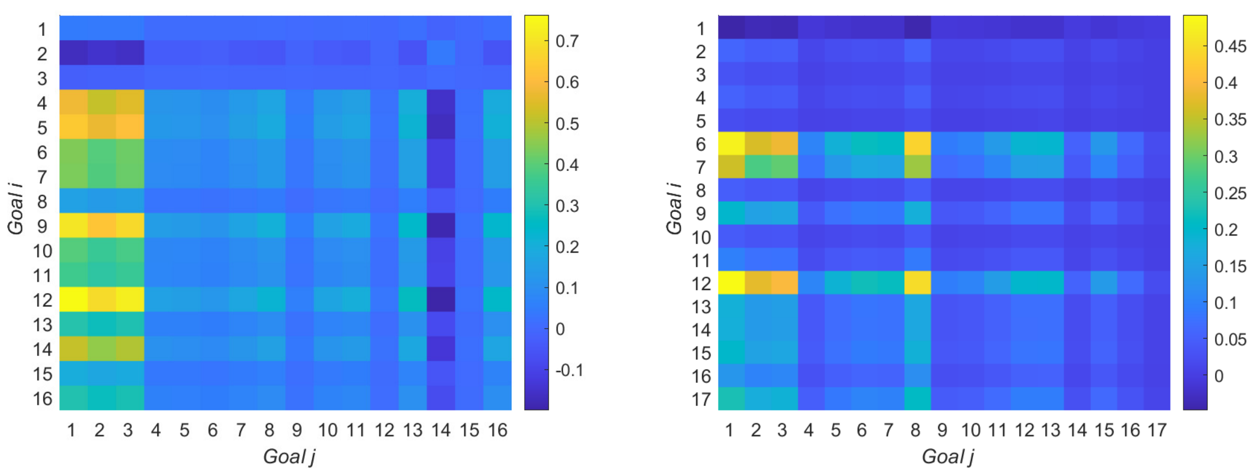
$$S_{ij}^{\mathrm{eq}} \quad := \quad \hat{\mathbf{n}} \cdot \frac{\partial \mathbf{v}^{(1)}}{\partial A_{ij}}, \tag{12}$$

that is, a 'scalar product' (sometimes referred to just as the 'dot product') between two vectors, one being the vector $\hat{\mathbf{n}} := \mathbb{1}/\sqrt{n}$ that describes complete equality of progress on all $n = 17$ SDGs, and the other being the rate of change in the leading eigenvector $\mathbf{v}^{(1)}$ when the $(i,j)^{\text{th}}$ element of A changes. The scalar product measures the extent to which the two vectors are pointing in the same direction, or not: if the vectors are at right-angles to each other then the scalar product is zero. The entries of the matrix S^{eq} therefore measure the extent to which a perturbation to A_{ij} increases the alignment of the leading eigenvector $\mathbf{v}^{(1)}$ with the vector $\hat{\mathbf{n}}$ which corresponds to complete equality of progress across all the SDGs. Mathematical details concerning (12) are given in Appendix 2 of Dawes (2022). The details are a little more complicated than for the growth rate sensitivity S^{m}_{ij} for unfortunately technical and purely mathematical reasons: because eigenvectors are defined only up to a normalisation, which we take to be the standard one that $|\mathbf{v}^{(1)}| = 1$, this normalisation has to be accounted for in the calculation of S^{eq}_{ij} so that an increase in one eigenvector component is compensated for by decreases elsewhere. For that reason we omit mathematical details here and refer the interested reader to Appendix B of Dawes (2022).

3.3 Illustrations of sensitivity analysis

Figures 6 and 7 illustrate the theory of the previous sections with calculations of the growth rate sensitivity S^{m} from (9) and the equality sensitivity S^{eq} from (12) for the ICSU and GSDR interlinkage matrices shown in figure 2.

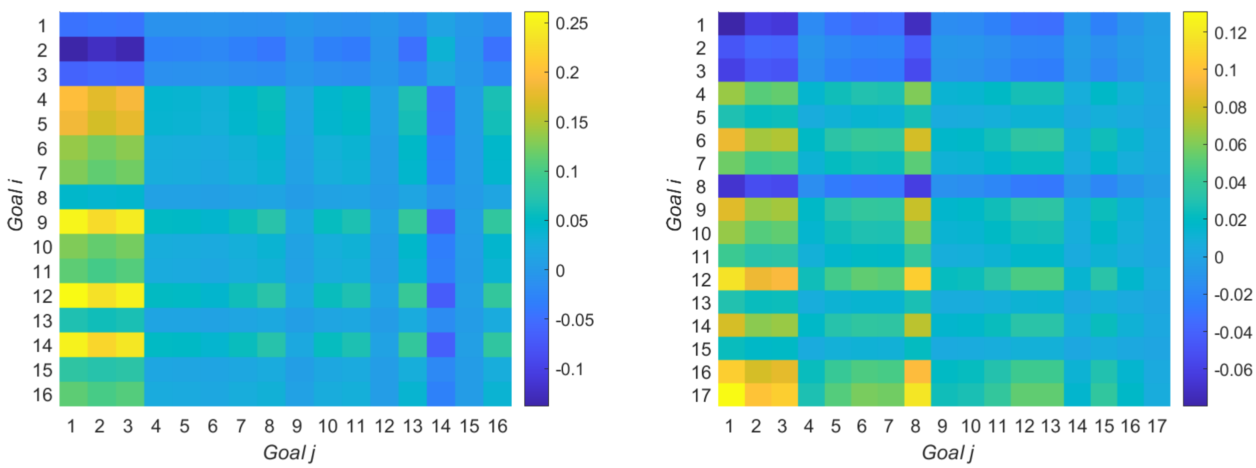
As we would anticipate from figures 3(b) and figure 4, the growth rate sensitivities are largest when both $\mathbf{y}^{(1)}$ and $\mathbf{v}^{(1)}$ take their most positive values, corresponding to changes in edges from nodes that are currently benefitting the most from the network structure (for example nodes 1-3 for the ICSU interlinkage matrix), to nodes that have the highest influences on other nodes (for example nodes 12, 9, 5 and 4 for the ICSU matrix). Hence increasing the values of edges $\{1,2,3\} \rightarrow \{4,5,9,12\}$ increases the overall growth rate of system-wide co-benefits the most, for the ICSU network. Similarly, for the GSDR network we would expect edges from $\{1,2,3,8\} \rightarrow \{6,7,12\}$ to affect the overall growth rate the most, and this is confirmed by figure 6(b).



(a) (b)

Figure 6: Growth rate sensitivities S^{m} for the leading eigenvalue λ_1 of (a) the ICSU network illustrated in figure 2(a); and (b) the averaged GSDR interaction network, illustrated in figure 2(b). Each square in the grid is coloured according to the relative increase in the growth rate (i.e. S^{m} that would result from a small increase in that specific interlinkage, keeping the remainder of the network the same. Lighter grey entries in the matrix indicate where increases in those interlinkages would increase S^{m} (and hence increase the leading eigenvalue λ_1); darker grey entries indicate where increases in an interlinkage would decrease S^{m} (and hence decrease the leading eigenvalue λ_1). Adapted under CC-BY-4.0 from figure 5 in Dawes (2022).

Figure 7 is a little harder to interpret, as the sensitivities plotted here reveal the extent to which the distribution of the co-benefits to different SDGs becomes more uniform as the weight of a specific edge is increased. But the general message is similar: increasing the interlinkages from SDGs 1-3 tends to re-balance the expected distribution of co-benefits across the SDGs, and this effect is particularly apparent when, for the ICSU network, the interlinkages are from $\{1,2,3\} \to \{4,5,9,12,14\}$. For the GSDR network, the edges that are most useful in this respect are those from $\{1,2,3,8\}$ to $\{12,17\}$.



(a)

(b)

Figure 7: Equality sensitivities S^{eq} for (a) the ICSU matrix and (b) the GSDR matrix. Light grey colours indicate that increasing those interlinkages serves to make the components of the leading eigenvector $\mathbf{v}^{(1)}$ more equal. Dark grey colours indicate that increasing an interlinkage will make the leading eigenvector more unbalanced. The relative differences between matrix entries are significant; the absolute differences between the colour scales in (a) and (b) are not significant. Adapted under CC-BY-4.0 from figure 6 in Dawes (2022).

4 Hierarchy

Building interlinkage matrices with *directed* edges between nodes, rather than *undirected* ones (which could be interpreted only as either mutually beneficial or mutually antagonistic relationships), immediately provokes questions at the system level as to whether the network taken as a whole has a sense of directionality to it. To a certain extent this has been present in the background to the discussions in section 2 since we have looked to identify nodes that are greater recipients of co-benefits, in some sense lying further 'downstream' of other nodes, and nodes that have the greatest influence and are providers of those benefits, in some sense lying further 'upstream' of other nodes. Thus there is implicitly a sense of directed order across the network. In this section we tackle this question of hierarchy directly.

It is clear that some directed networks do not have an overarching sense of hierarchy, for example consider the case of a network consisting of a cyclic ring of directed edges of equal edge weights. In such a network all nodes must lie at the same 'level' by symmetry. Contrastingly, a chain of positive directed edges all pointing in the same direction confers a clear hierarchy since, in our SDG context, progress on SDGs corresponding to nodes earlier in the chain results in co-benefits shared with SDGs further along the chain, but the reverse is not true. In the context of the SDGs it feels likely that there is at least some sense of hierarchy, for example it is clear that while progress on goals related to societal change such as SDG 5 (Gender equality) and SDG 10 (Reduced inequalities) are highly likely to lead to progress on 'human development' goals such as SDG 1 (No poverty) and SDG 3 (Good health and well-being), it is not so clear that all policy actions taken to reduce poverty or to improve healthcare would necessarily have co-benefits that included progress on SDG 5 or SDG

10; progress (at least to some extent) on SDG1 or SDG3 could be made through policy choices that fail to address persistent inequalities in gender or other population characteristics. Indeed, the persistent failure of sufficient data collection to monitor SDG 5 means that it may be difficult to quantify this effect although it surely exists (Pollitzer, 2022b).

In this section we look at a quantitative approach to deciding whether a given interlinkage network implies the existence of a hierarchy, and if it does, to computing the relative 'level' for different SDGs in the network.

4.1 Trophic confusion

Following Mackay et al (2020) we assign a level h_i to each node i (that is, each SDG in our goal-level analyses), and the challenge is to find a choice for the levels that minimizes a function F which measures the overall lack of directionality in the network (that is, we aim to choose levels in order to maximise the amount of directionality that exists). Mackay et al (2020) refer to F as the amount of 'trophic confusion' in the network, motivated by food webs where the structure of the trophic network is a key quantity of interest in order to understand an ecosystem. A generalization of their initial idea is to include a collection of pre-specified quantities g_{ij}, not necessarily related to the adjacency matrix A, that provide a set of target spacings between the levels h_i. The values of the levels h_i that minimize F can then be determined by minimizing, for example, the function

$$F(\boldsymbol{h}) := \frac{\sum_{ij} |A_{ij}| \, (h_i - h_j - g_{ij})^2}{\sum_{ij} |A_{ij}|} \, . \tag{13}$$

The form of $F(\boldsymbol{h})$ in (13) guarantees that it takes positive values, or is zero. F clearly has a minimum value zero that is achieved when the spacings between levels satisfy exactly the requirements given by the g_{ij}, if this is possible. If it is not, then minimising F attempts to get as close to satisfying this situation as possible. The denominator provides a normalization of the values of F, enabling comparisons between different networks if the distribution of edge weights are the same; otherwise it does not affect the values of the levels h_i that minimize F. Mathematically, by considering how F depends on the value chosen for a level h_i, it can be shown that the set of levels h_i that minimize F are equivalently given by solving the equation

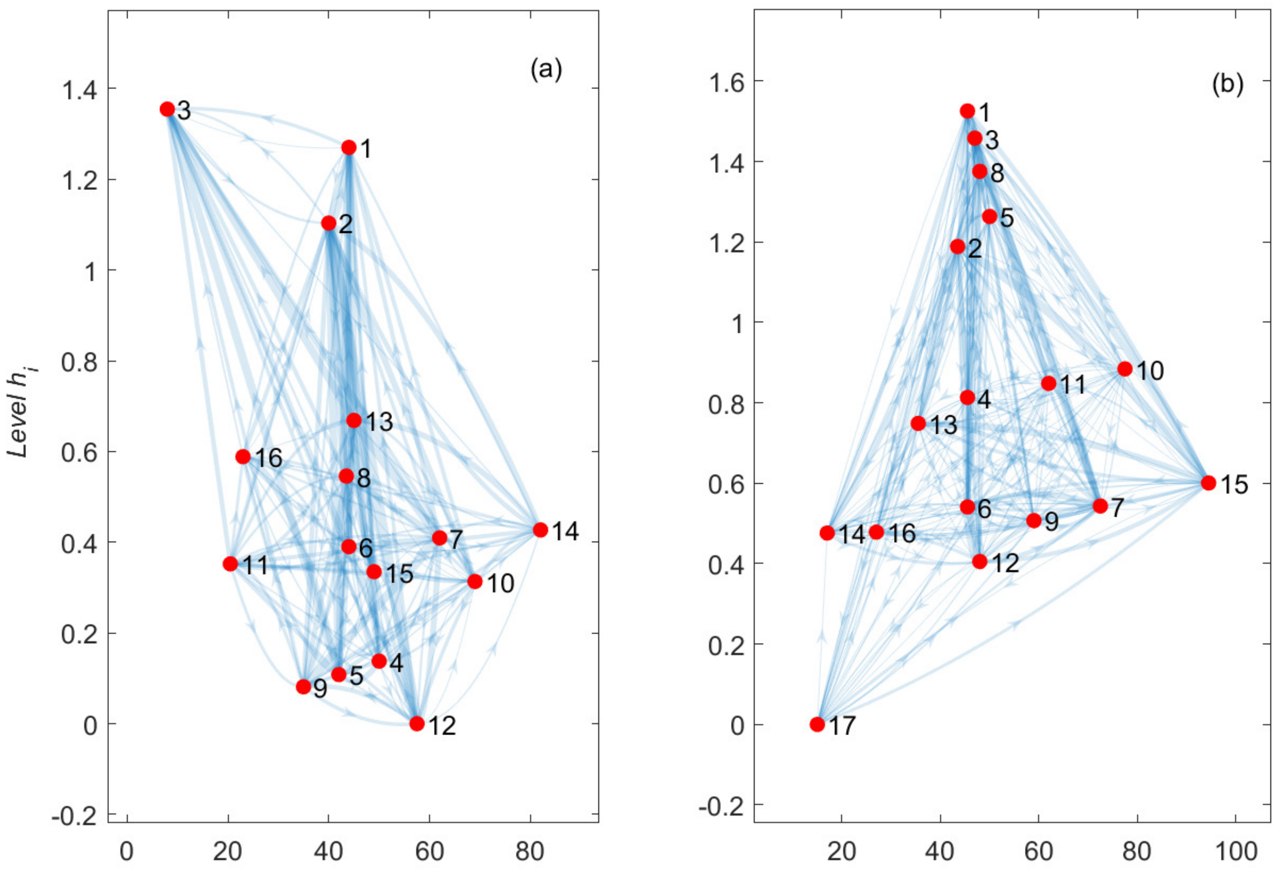
$$\Lambda \boldsymbol{h} \;\; = \;\; \mathrm{vecdiag}(AG^T - A^T G) \, , \tag{14}$$

where the matrix $\Lambda := \mathrm{diag}(\boldsymbol{k}) - A^T - A$, G is the matrix whose entries are given by the elements g_{ij}, and $\mathrm{vecdiag}(M)$ means the operation that extracts just the elements of the main diagonal of the matrix M and forms these elements into a column vector. Note also that $\boldsymbol{k}$ is the vector of total degrees $\boldsymbol{k} := \boldsymbol{k}^{\mathrm{in}} + \boldsymbol{k}^{\mathrm{out}}$ defined towards the start of section 2, G is the matrix whose entries are the target spacings g_{ij} and the notation $\mathrm{diag}(\mathbf{v})$ means the square matrix with the elements of the vector $\mathbf{v}$ as the main diagonal of the matrix, with zeros elsewhere. In the case $g_{ij} = 1$ for all i and j, the right-hand-side $\mathrm{vecdiag}(AG^T - A^T G)$ reduces to the difference between the (weighted) in-degree and out-degree.

Despite this rather detailed set of definitions, minimizing the 'trophic confusion' quantity F is straightforward to implement computationally and provides a representation of a network that as far as possible provides an overall sense of directionality. This enables a visualization of which nodes are furthest 'upstream' in the network and so influence most others, and which are furthest downstream and hence benefit most from co-benefits due to other SDGs.

4.2 Illustrations of hierarchy

In Figure 8 we show the results of minimizing the trophic confusion measure F defined in equation (13) for the ICSU and GSDR networks. In both parts of the figure, the relative levels of the SDGs are determined so that as many of the directed edges point upwards as possible. The relative horizontal positions of different SDGs have no meaning in terms of interlinkages; they are purely to make the figures as easy to interpret as possible. The influences between pairs of SDGs therefore generally run from the bottom of each figure to the top. We



(a) (b)

Figure 8: Hierarchies of nodes in the two networks. The vertical axes indicate the (relative) levels h_i, constructed so as to minimise the function F defined in equation (13), and placing the node at the lowest level at zero. The horizontal positions of nodes are chosen just so that as many links as possible in the network can be seen clearly. (a) Hierarchy of nodes in the ICSU matrix: SDG 12 (sustainable consumption and production) has most influence on other SDGs, whereas SDGs 1 – 3 are the most influenced by others; (b) Hierarchy of nodes in the GSDR matrix: SDG 17 (partnerships for the goals) is the most influencing, and SDGs 1, 3 and 8 are the most heavily influenced by the others. Adapted under CC-BY-4.0 from figure 10 in Dawes (2022).

see that consistently SDGs 1 – 3 appear close to the top of the figures, and SDG 12 appears low down. SDG 17 does not appear in Figure 8(a) since it was not included in the ICSU analysis. One surprise in Figure 8(b) is the relatively high position of SDG 5 (gender equality). This illustrates that the construction of these hierarchies depends on the robustness of the underlying data; as noted above, SDG 5 is rather sparsely represented in the literature on which the GSDR review was built. One useful direction for future research would clearly be to develop measures of the robustness or sensitivity of these hierarchy calculations to missing or suspected biases in the underlying data.

5 Summary and discussion

The aim of this chapter has been to introduce a number of concepts in network science and illustrate how they are able to contribute to the understanding of the study of SDG interlinkages. In particular they illuminate the challenge of drawing system-level inferences from a set of individual pairwise interactions between the 17 SDGs. Such system-level inferences would similarly be possible if we had started with data at the level of the 169 individual targets, or, indeed, at the even more granular level of indicator timeseries; the figures would have been more complicated to interpret but the methodology would apply unchanged. The extent to which SDG interlinkages reinforce each other, generating co-benefits, illustrates exactly the concept of policy coherence, with, naturally, some care required in the interpretation of these results depending on whether the interlinkage data are historic, related to current policy actions, or hypothesised as potentially existing if future policy actions are carried out.

In the public policy literature a central distinction is made between policy instruments (or tools, or techniques) and policy goals (outcomes). In terms of SDG interlinkages, the policy goals are made much more explicit than the policy instruments that might influence them; the mechanisms or inputs required to achieve the SDGs are often, as noted above, only described implicitly. The discussion of the relation between policy instruments and policy goals stretches back at least to the work of Jan Tinbergen who suggested that in typical situations, if n independent policy goals were to be achieved, then a set of at least n policy instruments would in general be required (Tinbergen, 1952; Schaeffer, 2019). But of course, any degree of alignment or coherence between policy goals might enable a reduction in the number of policy instruments required. A set of 17 separate policy instruments to achieve the SDGs would seem both awkward and at variance with the call for policy coherence, yet perhaps policy coherence is in many cases the atypical situation and needs to be 'designed in' to policy rather than being frequently and easily available.

The scope of the SDG agenda is so wide that the pursuit of some minimal number of policy 'tasks' that might generate policy coherence, or at least provide focal points for policy formulation, appears to be advocated by a number of recent proposals such as the Six Transformations proposed by the Sustainable Development Solutions Network (SDSN) (Sachs et al, 2019, 2020), or the six 'entry points' listed by the Global Sustainable Development Report 2019 (GSDR, 2019).

Three topics were discussed in this chapter: centrality in section 2, sensitivity in section 3, and hierarchy in section 4. Each topic was illustrated with results for two example networks; one based on expert review (the ICSU network, elaborated from the ICSU–ISSC report published in 2015) and the other on a substantial literature survey (GSDR, 2019). These were presented mainly in order to illustrate the mathematical tools but it is of interest to note the many similarities in the structure of the two networks built from these different data sources. It should be noted that the data sources are largely, but not quite fully, independent: one of the 177 assessments used by GSDR (2019) was ICSU-ISSC (2015). The results illustrated the kinds of overall conclusion that could be drawn about the structure of the interlinkage networks that was less obvious, and certainly less quantifiable, without carrying out the mathematical analysis set out above.

In section 2 the discussion focussed on 'eigencentrality'; in the context of directed networks, this corresponds to the idea that the most important nodes are those that receive the largest co-benefits from other nodes. As long as an interlinkage network does not contain too many negative entries (trade-offs), the interpretation of the eigenvector corresponding to the largest eigenvalue as a centrality score makes sense. It certainly continues to make sense as a measure of how much progress will be made on each SDG if the internal network dynamics and co-benefits or trade-offs within the network are left to dominate the dynamics of the system over time. Eigencentrality in the usual sense corresponds to the right eigenvector of the interlinkage matrix A. The left eigenvector of A has the dual interpretation: it shows which SDGs contribute the most to the co-benefits that are received by the SDGs that score highly in terms of eigencentrality (that is, those for which the corresponding entry in the right eigenvector is largest). For our illustrative examples, figure 3(b) and 4 reveal that SDGs 1–3 (and for the GSDR network, SDG 8 also) have high eigencentrality scores, and SDG 12 (and SDGs 6 and 7 for the GSDR network) have high influence scores.

Section 3 considered the robustness of the leading eigenvalue λ_1 of an interlinkage matrix (that is, the overall rate at which network effects influence progress) to perturbations in the interlinkage network. Two possible measures of this sensitivity were presented in equations (9) and 12 which describe either the sensitivity S^{m} of the growth rate λ_1 itself, or the sensitivity S^{eq} of the extent to which the network would result in equal progress across all SDGs, under perturbation. It turns out that S^{m} can be written elegantly as a product of the left and right eigenvectors of λ_1, but that S^{eq} is a little more complicated but yields similar results. These results show that adding interlinkages which allow goals that receive very high co-benefits to feed these back in and influence goals that received much lower levels of co-benefits are the ones that improve the overall rate of progress the most. While this is in line with our intuition, we have set out in this chapter a mathematical formulation that quantifies and justifies this statement. For our pair of example networks, we observe in figure 6 and 7 that the largest improvements in growth rates arise when progress on SDGs 1–3 (and SDG 8 for the GSDR network) are allowed to influence progress on specific other SDGs (including for both networks SDG 12).

The eigencentrality idea is particularly powerful due to its dynamical interpretation via equation (4); for an interlinkage matrix where there is a large gap between the leading eigenvalue and the remaining ones, eigencentrality reveals which SDGs will receive increasing shares of the overall co-benefits that are available. In cases where this gap is much smaller, more careful analysis may be required. There are of course many important additional caveats such as the need to consider interlinkages whose strength might vary over time, and the level of direct investment in support of specific SDGs. These issues are discussed further in Dawes (2022). The discussion in the present chapter is intended to highlight how one might begin to analyse a complex network of interactions and to move securely between consideration of individual interactions and the overall system-level consequences.

Finally, in section 4 we summarised another approach to this question of downstream response and upstream influence in the network, introducing a measure of hierarchy within the network, and computing the relative levels of different SDGs within the network (that is, how far upstream or downstream of other SDGs they are). Despite the differences in methodology, the results of the hierarchy calculations, shown in figure 8, appear in many examples to be well-correlated with the difference, for each SDG, between their scores $\mathbf{v}^{(1)}$ for receipt of co-benefits and $\mathbf{y}^{(1)}$ for generation of influence. This can be mathematically justified through careful consideration of (14); the solution for the levels $\boldsymbol{h}$ can be well-approximated by the formula

$$h_i \;\approx\; \frac{k_i^{\mathrm{in}} - k_i^{\mathrm{out}}}{k_i^{\mathrm{in}} + k_i^{\mathrm{out}}} \;\approx\; \frac{v_i^{(1)}\left(\mathbf{v}^{(1)} \cdot \mathbb{1}\right) - y_i^{(1)}\left(\mathbf{y}^{(1)} \cdot \mathbb{1}\right)}{v_i^{(1)}\left(\mathbf{v}^{(1)} \cdot \mathbb{1}\right) + y_i^{(1)}\left(\mathbf{y}^{(1)} \cdot \mathbb{1}\right)},$$

where the last expression on the right hand side is a further approximation in the limit in which the leading eigenvalue λ_1 is much larger than the remaining eigenvalues $\lambda_2, \ldots, \lambda_n$ in which case the 'network response' given by the eigencentrality measures $\mathbf{v}^{(1)}$ and $\mathbf{y}^{(1)}$ effectively captures the whole network response. Note that the coefficients in parentheses, for example $[\mathbf{v}^{(1)} \cdot \mathbb{1}]$, represent scalar ('dot') products of two vectors and are a convenient way to denote the sum of the entries in (in this example) the vector $\mathbf{v}^{(1)}$. The form of these coefficients is not particularly relevant for this discussion – of more interest is that this expression for h_i essentially looks like the difference $v_i^{(1)} - y_i^{(1)}$. In other words, the 'trophic levels' $\boldsymbol{h}$ are closely related to the (relative) differences between the influences received and the influences exerted, for each node in the network.

Of the many directions for future research we will touch on a couple of the most obvious and pressing. First, the degree to which the results are sensitive to biases, missing input data, and the details of the construction of any interlinkage network are obviously extremely important if the results are to be robust and valuable. In part, detailed knowledge of the limitations, or indeed the rationale behind any one interlinkage network should be known and understood at the time of construction, or data collection. The results for that network must then be interpreted in that light; without that context and background understanding interpretation of the results is likely always to be misleading. That is one reason to include the second topic, on sensitivity analysis, here.

It is likely that similar approaches can be developed in order to understand the robustness of the hierarchy calculations presented in section 4 as well. The hierarchy calculations can also be formulated in many different ways, for example making different choices for the target spacings g_{ij} in equation (13). There are indeed a number of different 'natural' choices for g_{ij}, and more generally g_{ij} allows the formulation of some kind of 'prior structure' that one might wish to impose on the answer, for example pairs of SDGs whose levels should be more or less closely aligned than other pairs of SDGs due to some structural influence or geographical restriction. These issues deserve to be the subject of future investigation. One concrete example of a set of targets that thematically cut across the SDGs are the 'access-related' targets listed by Pollitzer (2022b), see pages 18-19 and 38-39 therein. More prosaically, one could use the g_{ij} terms to link together targets within each goal, to explore how hierarchy results are affected by the potential for unanticipated spill-overs and co-benefits to exist between targets that lie within the same SDG.

Other centrality measures, building on eigencentrality, might also be useful in order to explore more fully the connected nature of the network; this motivated the inclusion of the short discussion of Katz centrality in section 2.4 and the discussion of the related work by Weitz et al (2018) that followed. There is related recent literature that allows the inclusion of three-way interactions between nodes in a directed network (Arrigo, Higham & Tudisco, 2020, and references therein); this may be another helpful direction for SDG interlinkage networks in particular due to the prevalence of 'nexus' thinking between, for example, water, energy and food (Weitz, Nilsson & Davis, 2014; Yillia, 2016; Fader et al, 2018; van Noordwijk et al, 2018; Putra, Pradhan & Kropp, 2020).

To conclude, while the construction of interlinkage networks is in itself a demanding and complex task, it appears to be gaining momentum as a way to visualize structure within the set of SDGs; it follows that appropriate network tools should be used to extract the system-level metrics and conclusions that follow from the interlinkage network. This second step, while of course inheriting biases present in the network data, is crucial in understanding in policy terms the emergent features that the interlinkage network represents and how best they should be addressed. Moreover, network science tools directly offer insights that should challenge and provoke policy coherence and so lead to better policy development.

References

Anderson, C.C., Denich, M., Warchold, A., Kropp, J.P. and Pradhan, P., A systems model of SDG target influence on the 2030 Agenda for Sustainable Development. *Sustainability Science* **17**, 1459-–1472 (2022)

Arrigo, F., Higham D.J., and Tudisco, F. A framework for second-order eigenvector centralities and clustering coefficients. *Proc. Roy. Soc.* A, **476**: 20190724 (2020)

Bali Swain R., and Ranganathan, S. Modeling interlinkages between sustainable development goals using network analysis. *World Development* **138**, 105136 (2021)

Bavelas, A. Communication patterns in task-oriented groups. *J. Acoustical Soc. America* **22**, 725–730 (1950)

Bennich, T., Weitz, N., and Carlsen, H. Deciphering the scientific literature on SDG interactions: a review and reading guide. *Science of the Total Environment* **728**, 138405 (2020)

Bianconi, G. *Multilayer Networks: Structure and Function.* Oxford University Press, Oxford, UK (2018)

Bonacich, P. Factoring and weighting approaches to clique identification. *J. Mathematical Sociology* **2**, 113–120 (1972)

Bonacich, P. Some unique properties of eigenvector centrality. *Social Networks* **29**, 555-564 (2007)

Dawes, J.H.P. Are the Sustainable Development Goals self-consistent and mutually achievable? *Sustainable Development* **28**, 101–117 (2020)

Dawes, J.H.P. SDG interlinkage networks: analysis, robustness, sensitivities, and hierarchies. *World Development* **149**, 105693 (2022)

Dawes, J.H.P., Zhou, X. and Moinuddin, M. System-level consequences of synergies and trade-offs between SDGs: quantitative analysis of interlinkage networks at country level. *Sustainability Science* **17**, 1435—1457 ((2022)

Fader, M., Cranmer, C., Lawford, R., and Engel–Cox, J. Toward an understanding of synergies and trade-offs between water, energy, and food SDG targets. *Front. Environ. Sci.* **6**, 112 (2018)

Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry*, **40**(1), 35–41 (1977)

Independent Group of Scientists appointed by the Secretary-General, *Global Sustainable Development Report 2019: The Future is Now – Science for Achieving Sustainable Development*. United Nations: New York, USA (2019)

Greenbaum, A., Li, R.-C., and Overton, M. L. First-order perturbation theory for eigenvalues and eigenvectors. *SIAM Review* **62**(2), 463–482 (2020)

International Council for Science (ICSU); International Social Sciences Council (ISSC). *Review of the Sustainable Development Goals: The Science Perspective*. International Council for Science: Paris, France (2015)

Jacobi, C. J. G. Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen. *J. Reine und Angewandte Mathematik* **30**, 51–95 (1846)

Katz, L. A new status index derived from sociometric analysis. *Psychometrika* **18**, 39–43 (1953)

Landau, E. Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach* 11: 366–369 (1895)

Le Blanc, D. Towards integration at last? The sustainable development goals as a network of targets. UN Department of Economic & Social Affairs, DESA Working Paper No. 141 (2015)

Leontief, W.W. *The Structure of American Economy 1919-1929*. Harvard University Press, New Haven, CT. (1941)

MacKay, R.S., Johnson, S., Sansom B. How directed is a directed network? *Royal Society Open Science* **7**, 201138 (2020)

Meyer, C.D., *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2000)

Neubert, M. G., and Caswell, H. Alternatives to resilience for measuring the responses of ecological systems to perturbations. *Ecology* **78**(3), 653–665 (1997)

Newman, M. Networks: an introduction. 2nd edition. Oxford University Press: Oxford, UK (2018)

Nilsson, M., Griggs, D., and Visbeck, M. Policy: map the interactions between sustainable development goals. *Nature*, 534, 320–322 (2016)

Pedercini, M., Arquitt, S., Collste, D. and Herren, H. Harvesting synergy from sustainable development goal interactions *Proc. Nat. Acad. Sci.* **116**(46) 23021–23028 (2019)

Pham-Truffert, M., Metz, F., Fischer, M., Rueff, H., and Messerli, P. Interactions among Sustainable Development Goals: Knowledge for identifying multipliers and virtuous cycles. *Sustainable Development* 1–15 (2020) https://doi.org/10.1002/sd.2073

Pham-Truffert, M., Rueff, H., and Messerli, P. Knowledge for Sustainable Development: Interactive repository of SDG interactions. CDEdatablog, Bern, Switzerland: CDE. https://datablog.cde.unibe.ch/index.php/2019/08/29/sdg-interactions/ (2019). Accessed on 7 September 2020.

Pollitzer, E. *Integrating a gender perspective into implementation of Agenda 2030*. Guidance Note. UN Women: Nairobi, Kenya (2022)

Pollitzer, E. *Gender analysis toolkit for prioritising SDG goals and targets: How to improve the knowledge base for SDGs, remove gender gaps in evidence, and integrate gender perspectives into analyses of SDGs.* UN Women: Nairobi, Kenya (2022)

Pradhan, P., Costa, L., Rybski, D., Lucht, W., and Kropp, J.P. A Systematic Study of Sustainable Development Goal (SDG) Interactions. *Earth's Future*, **5**, 1169-–1179 (2017)

Putra, M.P.I.F., Pradhan, P., and Kropp, J.P. A systematic analysis of Water-Energy-Food security nexus: a South Asian case study. *Science of the Total Environment* **728**, 138451 (2020)

Sachs, J., Schmidt–Traub, G., Kroll, C., Lafortune, G. and Fuller, G. *Sustainable Development Report 2019: G20 and Large Countries edition*. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN). (2019)

Sachs, J., Schmidt–Traub, G., Kroll, C., Lafortune, G., Fuller, G. and Woelm, F. The Sustainable Development Goals and COVID-19. Sustainable Development Report 2020. Cambridge: Cambridge University Press (2020)

Schaeffer, P.V., and Willardsen, K. *A Note on the Tinbergen Rule*. Working paper. Accessed in July 2021 from `https://www.petervschaeffer.com/publications.html`. (2019)

Seeley, J.R. The net of reciprocal influence: A problem in treating sociometric data. *The Canadian Journal of Psychology* **3**(4), 234—240 (1949)

Tinbergen, J. *On the Theory of Economic Policy*. North–Holland, New York (1952)

United Nations, *Transforming our world: The 2030 Agenda for Sustainable Development*. A/RES/70/1, United Nations General Assembly, New York (2015)

van Noordwijk, M., Duguma, L.A., Dewi, S., Leimona, B., Catacutan, D.C., Lusiana, B., Öborn, I., Hairiah, K., and Minang, P.A. SDG synergy between agriculture and forestry in the food, energy, water and income nexus: reinventing agroforestry? *Current opinion in environmental sustainability*, **34**, 33–42 (2018)

van Soest, H.L., van Vuuren, D.P., Hilaire, J., Minx, J.C., Harmsen, M.J.H.M., Krey, V, Popp, A., Riahi, K., and Luderer, G. Analysing interactions among Sustainable Development Goals with Integrated Assessment Models. *Global Transitions* **1**, 210–225 (2019)

Wasserman, S. and Faust, K. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, UK (1994)

Weitz, N., Carlsen, H., Nilsson, M., and Skånberg, K. Towards systemic and contextual priority setting for implementing the 2030 Agenda. *Sustain. Sci.* **13**, 531-548 (2018)

Weitz, N., Nilsson, M., and Davis, M. A nexus approach to the post-2015 agenda: formulating integrated water, energy, and food SDGs. *SAIS Review of International Affairs*, **34** (2), 37–50 (2014)

Yillia, P.T. Water-energy-food nexus: framing the opportunities, challenges and synergies for implementing the SDGs. *Österr. Wasser–Abfallwirtsch*. **68**, 86–98 (2016)

Zhou, X. and Moinuddin, M. Sustainable Development Goals Interlinkages and Network Analysis: A practical tool for SDG integration and policy coherence. Research Report. IGES (2017)

Zhou, X., Moinuddin, M., and Li, Y. SDG Interlinkages Analysis & Visualisation Tool (V3.0) `https://sdginterlinkages.iges.jp/index.html` (2019)