

Predicting metawebs: graph embeddings can help alleviate spatial data deficiencies

Tanya Strydom^{1,2,‡} Timothée Poisot^{1,2,‡}

¹ Département de Sciences Biologiques, Université de Montréal, Montréal, Canada ² Quebec Centre for Biodiversity Science, Montréal, Canada

[‡] These authors contributed equally to the work

Correspondance to:

Timothée Poisot — timothee.poisot@umontreal.ca

1. Metawebs, i.e. networks of potential interactions within a species pool, are a powerful abstraction to understand how large-scales species interaction networks are structured.
2. Because metawebs are typically expressed at large spatial and taxonomic scales, assembling them is a tedious and costly process; predictive methods can help circumvent the limitations in data deficiencies, by providing 'draft' metawebs.
3. One way to improve the predictive ability is to maximize the information used for prediction, by using graph embeddings rather than the list of species interactions. Graph embedding is an emerging field in machine learning that holds great potential for ecological problems.
4. In this perspective, we outline how the challenges associated with inferring metawebs line-up with the advantages of graph embeddings; furthermore, because metawebs are inherently spatial objects, we discuss how the choice of the species pool has consequences on the reconstructed network, but also embeds hypotheses about which human-made boundaries are ecologically meaningful.

1 Having a general solution for inferring *potential* interactions (despite the unavailability of interaction data)
2 could be the catalyst for significant breakthroughs in our ability to start thinking about species interaction
3 networks over large spatial scales (Hortal et al., 2015). In a recent overview of the field of ecological
4 network prediction, Strydom et al. (2021) identified two challenges of interest to the prediction of
5 interactions at large scales. First, there is a relative scarcity of relevant data in most places globally –
6 paradoxically, this restricts our ability to infer interactions to locations where inference is perhaps the least
7 required; second, accurate predictions often demand accurate predictors, and the lack of methods that can
8 leverage small amount of data is a serious impediment to our predictive ability globally.

9 Following the definition of Dunne (2006), a metaweb is a network analogue to the regional species pool;
10 specifically, it is an inventory of all *potential* interactions between species from a spatially delimited area
11 (and so captures the γ diversity of interactions). The metaweb is, therefore, *not* a prediction of the food
12 web at a specific locale within the spatial area it covers, and will have a different structure (notably by
13 having a larger connectance; see e.g. Wood et al., 2015). These local food webs (which captures the α
14 diversity of interactions) are a subset of the metaweb’s species and interactions, and have been called
15 “metaweb realizations” (Poisot et al., 2015). Differences between local food web and their metaweb are
16 due to chance, species abundance and co-occurrence, local environmental conditions, and local
17 distribution of functional traits, among others.

18 Because the metaweb represents the joint effect of functional, phylogenetic, and macroecological
19 processes (Morales-Castilla et al., 2015), it holds valuable ecological information. Specifically, it is the
20 “upper bounds” on what the composition of the local networks can be (see e.g. McLeod et al., 2021). These
21 local networks, in turn, can be reconstructed given appropriate knowledge of local species composition,
22 providing information on structure of food webs at finer spatial scales. This has been done for example for
23 tree-galler-parasitoid systems (Gravel et al., 2018), fish trophic interactions (Albouy et al., 2019), tetrapod
24 trophic interactions (O’Connor et al., 2020), and crop-pest networks (Grünig et al., 2020). Whereas the
25 original metaweb definition, and indeed most past uses of metawebs, was based on the presence/absence
26 of interactions, we focus on *probabilistic* metawebs where interactions are represented as the chance of
27 success of a Bernoulli trial (see e.g. Poisot et al., 2016); therefore, not only does our method recommend
28 interactions that may exist, it gives each interaction a score, allowing us to properly weigh them.

29 **The metaweb is an inherently probabilistic object**

30 Yet, owing to the inherent plasticity of interactions, there have been documented instances of food webs
31 undergoing rapid collapse/recovery cycles over short periods of time (Pedersen et al., 2017). The
32 embedding of a network, in a sense, embeds its macro-evolutionary history, especially as RDPG captures
33 ecological signal (Dalla Riva & Stouffer, 2016); at this point, it is important to recall that a metaweb is
34 intended as a catalogue of all potential interactions, which should then be filtered (Morales-Castilla et al.,
35 2015). In practice (and in this instance) the reconstructed metaweb will predict interactions that are
36 plausible based on the species' evolutionary history, however some interactions would/would not be
37 realized due to human impact.

38 Dallas et al. (2017) suggested that most links in ecological networks may be cryptic, *i.e.* uncommon or
39 otherwise hard to observe. This argument essentially echoes Jordano (2016): the sampling of ecological
40 interactions is difficult because it requires first the joint observation of two species, and then the
41 observation of their interaction. In addition, it is generally expected that weak or rare links would be more
42 common in networks (Csermely, 2004), compared to strong, persistent links; this is notably the case in
43 food chains, wherein many weaker links are key to the stability of a system (Neutel et al., 2002). In the
44 light of these observations, the results in fig. ?? are not particularly surprising: we expect to see a surge in
45 these low-probability interactions under a model that has a good predictive accuracy. Because the
46 predictions we generate are by design probabilistic, then one can weigh these rare links appropriately. In a
47 sense, that most ecological interactions are elusive can call for a slightly different approach to sampling:
48 once the common interactions are documented, the effort required in documenting each rare interaction
49 may increase exponentially. Recent proposals suggest that machine learning algorithms, in these
50 situations, can act as data generators (Hoffmann et al., 2019): in this perspective, high quality
51 observational data can be supplemented with synthetic data coming from predictive models, which
52 increases the volume of information available for inference. Indeed, Strydom et al. (2021) suggested that
53 knowing the metaweb may render the prediction of local networks easier, because it fixes an “upper
54 bound” on which interactions can exist; indeed, with a probabilistic metaweb, we can consider that the
55 metaweb represents an aggregation of informative priors on the interactions.

56 Graph embedding offers promises for the inference of potential 57 interactions

58 Graph embedding is a varied family of machine learning techniques aiming to transform nodes and edges
59 into a vector space, usually of a lower dimension, whilst maximally retaining key properties of the graph
60 (Yan et al., 2005). Ecological networks are an interesting candidate for the widespread application of
61 embeddings, as they tend to possess a shared structural backbone (Mora et al., 2018), which hints at
62 structural invariants that can be revealed at lower dimensions. Indeed, previous work by Eklöf et al. (2013)
63 suggests that food webs are inherently low-dimensional objects, and can be adequately represented with
64 less than ten dimensions. Simulation results by Botella et al. (2022) suggest that there is no best method to
65 identify architectural similarities between networks, and that multiple approaches need to be tested and
66 compared to the network descriptor of interest.

67 But the popularity of graph embedding techniques in machine learning is rather more intuitive than the
68 search for structural invariants: while graphs are discrete objects, machine learning techniques tend to
69 handle continuous data better. Therefore, bringing a discrete graph into a continuous vector space opens
70 up a broader variety of predictive algorithms.

Table 1: Overview of some common graph embedding approaches, by time of publication, alongside examples of their use in the prediction of species interactions. Surprisingly, these methods have not yet been used routinely to predict species interactions; most of the examples we identified were either statistical associations, or analogues to joint species distribution models. ^a: statistical interactions; ^b: joint-SDM-like approach.

Method	Embedding approach	Reference	Species interactions application
RDPG	graphs through SVD	Young & Scheinerman (2007)	Poisot et al. (2021)
tSNE	nodes through statistical divergence	Hinton & Roweis (2002)	Cieslak et al. (2020) ^a
DeepWalk	graph walk	Perozzi et al. (2014)	Wardeh et al. (2021)

Method	Embedding approach	Reference	Species interactions application
FastEmbed	graph through PCA/SVD analogue	Ramasamy & Madhow (2015)	
LINE	nodes through statistical divergence	Tang et al. (2015)	
SDNE	nodes through auto-encoding	D. Wang et al. (2016)	
node2vec	node embedding	Grover & Leskovec (2016)	
graph2vec	sub-graph embedding	Narayanan et al. (2017)	
DMSE	joint node embedding	D. Chen et al. (2017)	D. Chen et al. (2017) ^b
HARP	nodes through a meta-strategy	H. Chen et al. (2017)	
GraphKKE	graph embedding	Melnyk et al. (2020)	Melnyk et al. (2020) ^a
Joint methods	multiple graphs	S. Wang et al. (2021)	

71 **The metaweb embeds hypotheses about which spatial boundaries are**
72 **meaningful**

73 As Herbert (1965) rightfully pointed out, “[y]ou can’t draw neat lines around planet-wide problems”; in
74 this regard, our approach (and indeed, any inference of a metaweb at large scales) must contend with
75 several interesting and interwoven families of problems. The first is the limit of the metaweb to embed
76 and transfer. If the initial metaweb is too narrow in scope, notably from a taxonomic point of view, the
77 chances of finding another area with enough related species to make a reliable inference decreases; this

would likely be indicated by large confidence intervals during ancestral character estimation, but the lack of well documented metawebs is currently preventing the development of more concrete guidelines. The question of phylogenetic relatedness and dispersal is notably true if the metaweb is assembled in an area with mostly endemic species, and as with every predictive algorithm, there is room for the application of our best ecological judgement. Conversely, the metaweb should be reliably filled, which assumes that the S^2 interactions in a pool of S species have been examined, either through literature surveys or expert elicitation. Supp. Mat. 1 provides some guidance as to the type of sampling effort that should be prioritized. While RDPG was able to maintain very high predictive power when interactions were missing, the addition of false positive interactions was immediately detected; this suggests that it may be appropriate to err on the side of “too many” interactions when constructing the initial metaweb to be transferred. The second series of problems are related to determining which area should be used to infer the new metaweb in, as this determines the species pool that must be used.

In our application, we focused on the mammals of Canada. The upside of this approach is that information at the country level is likely to be required by policy makers and stakeholders for their biodiversity assessment, as each country tends to set goals at the national level (Buxton et al., 2021) for which quantitative instruments are designed (Turak et al., 2017), with specific strategies often enacted at smaller scales (Ray et al., 2021). And yet, we do not really have a satisfying answer to the question of “where does a food web stop?”; the current most satisfying solutions involve examining the spatial consistency of network area relationships (see e.g. Galiana et al., 2018, 2019, 2021; **Fortin2021NetEco?**), which is of course impossible in the absence of enough information about the network itself. This suggests that an *a posteriori* refinement of the results may be required, based on a downscaling of the metaweb. The final family of problems relates less to the availability of data or quantitative tools, and more to the praxis of spatial ecology. Operating under the context of national divisions, in large parts of the world, reflects nothing more than the legacy of settler colonialism. Indeed, the use of ecological data is not an apolitical act (Nost & Goldstein, 2021), as data infrastructures tend to be designed to answer questions within national boundaries, and their use both draws upon and reinforces territorial statecraft; as per Machen & Nost (2021), this is particularly true when the output of “algorithmic thinking” (e.g. relying on machine learning to generate knowledge) can be re-used for governance (e.g. enacting conservation decisions at the national scale). We therefore recognize that methods such as we propose operate under the framework that contributed to the ongoing biodiversity crisis (Adam, 2014), reinforced

108 environmental injustice (Choudry, 2013; Domínguez & Luoma, 2020), and on Turtle Island especially,
109 should be replaced by Indigenous principles of land management (Eichhorn et al., 2019; No'kmaq et al.,
110 2021). As we see AI/ML being increasingly mobilized to generate knowledge that is lacking for
111 conservation decisions (e.g. Lamba et al., 2019; Mosebo Fernandes et al., 2020), our discussion of these
112 tools need to go beyond the technical, and into the governance consequences they can have.

113 **Acknowledgements:** We acknowledge that this study was conducted on land within the traditional
114 unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and
115 Omàmiwininiwak nations. TP, TS, DC, and LP received funding from the Canadian Institute for Ecology &
116 Evolution. FB is funded by the Institute for Data Valorization (IVADO). TS, SB, and TP are funded by a
117 donation from the Courtois Foundation. CB was awarded a Mitacs Elevate Fellowship no. IT12391, in
118 partnership with fRI Research, and also acknowledges funding from Alberta Innovates and the Forest
119 Resources Improvement Association of Alberta. M-JF acknowledges funding from NSERC Discovery
120 Grant and NSERC CRC. RR is funded by New Zealand's Biological Heritage Ngā Koiora Tuku Iho
121 National Science Challenge, administered by New Zealand Ministry of Business, Innovation, and
122 Employment. BM is funded by the NSERC Alexander Graham Bell Canada Graduate Scholarship and the
123 FRQNT master's scholarship. LP acknowledges funding from NSERC Discovery Grant (NSERC
124 RGPIN-2019-05771). TP acknowledges financial support from NSERC through the Discovery Grants and
125 Discovery Accelerator Supplement programs.

126 References

- 127 Adam, R. (2014). *Elephant treaties: The Colonial legacy of the biodiversity crisis*. UPNE.
- 128 Albouy, C., Archambault, P., Appeltans, W., Araújo, M. B., Beauchesne, D., Cazelles, K., Cirtwill, A. R.,
129 Fortin, M.-J., Galiana, N., Leroux, S. J., Pellissier, L., Poisot, T., Stouffer, D. B., Wood, S. A., & Gravel, D.
130 (2019). The marine fish food web is globally connected. *Nature Ecology & Evolution*, 3(8, 8),
131 1153–1161. <https://doi.org/10.1038/s41559-019-0950-y>
- 132 Botella, C., Dray, S., Matias, C., Miele, V., & Thuiller, W. (2022). An appraisal of graph embeddings for
133 comparing trophic network architectures. *Methods in Ecology and Evolution*, 13(1), 203–216.
134 <https://doi.org/10.1111/2041-210X.13738>

135 Buxton, R. T., Bennett, J. R., Reid, A. J., Shulman, C., Cooke, S. J., Francis, C. M., Nyboer, E. A., Pritchard,
 136 G., Binley, A. D., Avery-Gomm, S., Ban, N. C., Beazley, K. F., Bennett, E., Blight, L. K., Bortolotti, L. E.,
 137 Camfield, A. F., Gadallah, F., Jacob, A. L., Naujokaitis-Lewis, I., ... Smith, P. A. (2021). Key
 138 information needs to move from knowledge to action for biodiversity conservation in Canada.
 139 *Biological Conservation*, 256, 108983. <https://doi.org/10.1016/j.biocon.2021.108983>

140 Chen, D., Xue, Y., Fink, D., Chen, S., & Gomes, C. P. (2017). *Deep Multi-species Embedding*. 3639–3646.
 141 <https://www.ijcai.org/proceedings/2017/509>

142 Chen, H., Perozzi, B., Hu, Y., & Skiena, S. (2017, November 16). *HARP: Hierarchical Representation*
 143 *Learning for Networks*. <http://arxiv.org/abs/1706.07845>

144 Choudry, A. (2013). Saving biodiversity, for whom and for what? Conservation NGOs, complicity,
 145 colonialism and conquest in an era of capitalist globalization. In *NGOization: Complicity,*
 146 *contradictions and prospects* (pp. 24–44). Bloomsbury Publishing.

147 Cieslak, M. C., Castelfranco, A. M., Roncalli, V., Lenz, P. H., & Hartline, D. K. (2020). T-Distributed
 148 Stochastic Neighbor Embedding (t-SNE): A tool for eco-physiological transcriptomic analysis. *Marine*
 149 *Genomics*, 51, 100723. <https://doi.org/10.1016/j.margen.2019.100723>

150 Csermely, P. (2004). Strong links are important, but weak links stabilize them. *Trends in Biochemical*
 151 *Sciences*, 29(7), 331–334. <https://doi.org/10.1016/j.tibs.2004.05.004>

152 Dalla Riva, G. V., & Stouffer, D. B. (2016). Exploring the evolutionary signature of food webs' backbones
 153 using functional traits. *Oikos*, 125(4), 446–456. <https://doi.org/10.1111/oik.02305>

154 Dallas, T., Park, A. W., & Drake, J. M. (2017). Predicting cryptic links in host-parasite networks. *PLOS*
 155 *Computational Biology*, 13(5), e1005557. <https://doi.org/10.1371/journal.pcbi.1005557>

156 Domínguez, L., & Luoma, C. (2020). Decolonising Conservation Policy: How Colonial Land and
 157 Conservation Ideologies Persist and Perpetuate Indigenous Injustices at the Expense of the
 158 Environment. *Land*, 9(3, 3), 65. <https://doi.org/10.3390/land9030065>

159 Dunne, J. A. (2006). The Network Structure of Food Webs. In J. A. Dunne & M. Pascual (Eds.), *Ecological*
 160 *networks: Linking structure and dynamics* (pp. 27–86). Oxford University Press.

161 Eichhorn, M. P., Baker, K., & Griffiths, M. (2019). Steps towards decolonising biogeography. *Frontiers of*
 162 *Biogeography*, 12(1), 1–7. <https://doi.org/10.21425/F5FBG44795>

163 Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N. P., Dalsgaard, B., de Sassi, C., Galetti,
 164 M., Guimarães, P. R., Lomáscolo, S. B., Martín González, A. M., Pizo, M. A., Rader, R., Rodrigo, A.,
 165 Tylíanakis, J. M., Vázquez, D. P., & Allesina, S. (2013). The dimensionality of ecological networks.
 166 *Ecology Letters*, 16(5), 577–583. <https://doi.org/10.1111/ele.12081>

167 Galiana, N., Barros, C., Braga, J., Ficetola, G. F., Maiorano, L., Thuiller, W., Montoya, J. M., & Lurgi, M.
 168 (2021). The spatial scaling of food web structure across European biogeographical regions. *Ecography*,
 169 *n/a*(*n/a*). <https://doi.org/10.1111/ecog.05229>

170 Galiana, N., Hawkins, B. A., & Montoya, J. M. (2019). The geographical variation of network structure is
 171 scale dependent: Understanding the biotic specialization of host–parasitoid networks. *Ecography*,
 172 42(6), 1175–1187. <https://doi.org/10.1111/ecog.03684>

173 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., Gravel, D., & Montoya,
 174 J. M. (2018). The spatial scaling of species interaction networks. *Nature Ecology & Evolution*, 2(5),
 175 782–790. <https://doi.org/10.1038/s41559-018-0517-3>

176 Gravel, D., Baiser, B., Dunne, J. A., Kopelke, J.-P., Martinez, N. D., Nyman, T., Poisot, T., Stouffer, D. B.,
 177 Tylíanakis, J. M., Wood, S. A., & Roslin, T. (2018). Bringing Elton and Grinnell together: A quantitative
 178 framework to represent the biogeography of ecological interaction networks. *Ecography*, 0(0).
 179 <https://doi.org/10.1111/ecog.04006>

180 Grover, A., & Leskovec, J. (2016). Node2vec: Scalable Feature Learning for Networks. *Proceedings of the*
 181 *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 855–864.
 182 <https://doi.org/10.1145/2939672.2939754>

183 Grünig, M., Mazzi, D., Calanca, P., Karger, D. N., & Pellissier, L. (2020). Crop and forest pest metawebs
 184 shift towards increased linkage and suitability overlap under climate change. *Communications Biology*,
 185 3(1, 1), 1–10. <https://doi.org/10.1038/s42003-020-0962-9>

186 Herbert, F. (1965). *Dune* (1st ed.). Chilton Book Company.

187 Hinton, G., & Roweis, S. T. (2002). Stochastic neighbor embedding. *NIPS*, 15, 833–840.

188 Hoffmann, J., Bar-Sinai, Y., Lee, L. M., Andrejevic, J., Mishra, S., Rubinstein, S. M., & Rycroft, C. H. (2019).
 189 Machine learning in a data-limited regime: Augmenting experiments with synthetic data uncovers
 190 order in crumpled sheets. *Science Advances*, 5(4), eaau6792.

191 <https://doi.org/10.1126/sciadv.aau6792>

192 Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven
 193 Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and*
 194 *Systematics*, 46(1), 523–549. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>

195 Jordano, P. (2016). Sampling networks of ecological interactions. *Functional Ecology*, 30(12), 1883–1893.
 196 <https://doi.org/10.1111/1365-2435.12763>

197 Lamba, A., Cassey, P., Segaran, R. R., & Koh, L. P. (2019). Deep learning for environmental conservation.
 198 *Current Biology*, 29(19), R977–R982. <https://doi.org/10.1016/j.cub.2019.08.016>

199 Machen, R., & Nost, E. (2021). Thinking algorithmically: The making of hegemonic knowledge in climate
 200 governance. *Transactions of the Institute of British Geographers*, 46(3), 555–569.
 201 <https://doi.org/10.1111/tran.12441>

202 McLeod, A., Leroux, S. J., Gravel, D., Chu, C., Cirtwill, A. R., Fortin, M.-J., Galiana, N., Poisot, T., & Wood,
 203 S. A. (2021). Sampling and asymptotic network properties of spatial multi-trophic networks. *Oikos*,
 204 n/a(n/a). <https://doi.org/10.1111/oik.08650>

205 Melnyk, K., Klus, S., Montavon, G., & Conrad, T. O. F. (2020). GraphKKE: Graph Kernel Koopman
 206 embedding for human microbiome analysis. *Applied Network Science*, 5(1), 96.
 207 <https://doi.org/10.1007/s41109-020-00339-2>

208 Mora, B. B., Gravel, D., Gilarranz, L. J., Poisot, T., & Stouffer, D. B. (2018). Identifying a common backbone
 209 of interactions underlying food webs from different ecosystems. *Nature Communications*, 9(1), 2603.
 210 <https://doi.org/10.1038/s41467-018-05056-0>

211 Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from
 212 proxies. *Trends in Ecology & Evolution*, 30(6), 347–356.
 213 <https://doi.org/10.1016/j.tree.2015.03.014>

214 Mosebo Fernandes, A. C., Quintero Gonzalez, R., Lenihan-Clarke, M. A., Leslie Trotter, E. F., & Jokar
 215 Arsanjani, J. (2020). Machine Learning for Conservation Planning in a Changing Climate.
 216 *Sustainability*, 12(18, 18), 7657. <https://doi.org/10.3390/su12187657>

217 Narayanan, A., Chandramohan, M., Venkatesan, R., Chen, L., Liu, Y., & Jaiswal, S. (2017, July 17).
 218 *Graph2vec: Learning Distributed Representations of Graphs*. <http://arxiv.org/abs/1707.05005>

219 Neutel, A.-M., Heesterbeek, J. A. P., & de Ruiter, P. C. (2002). Stability in Real Food Webs: Weak Links in
 220 Long Loops. *Science*, 296(5570), 1120–1123. <https://doi.org/10.1126/science.1068326>

221 No'kmaq, M., Marshall, A., Beazley, K. F., Hum, J., Joudry, Shalan, Papadopoulos, A., Pictou, S., Rabesca,
 222 J., Young, L., & Zurba, M. (2021). “Awakening the sleeping giant”: Re-Indigenization principles for
 223 transforming biodiversity conservation in Canada and beyond. *FACETS*, 6(1), 839–869.

224 Nost, E., & Goldstein, J. E. (2021). A political ecology of data. *Environment and Planning E: Nature and*
 225 *Space*, 25148486211043503. <https://doi.org/10.1177/25148486211043503>

226 O'Connor, L. M. J., Pollock, L. J., Braga, J., Ficetola, G. F., Maiorano, L., Martinez-Almoyna, C.,
 227 Montemaggiore, A., Ohlmann, M., & Thuiller, W. (2020). Unveiling the food webs of tetrapods across
 228 Europe through the prism of the Eltonian niche. *Journal of Biogeography*, 47(1), 181–192.
 229 <https://doi.org/10.1111/jbi.13773>

230 Pedersen, E. J., Thompson, P. L., Ball, R. A., Fortin, M.-J., Gouhier, T. C., Link, H., Moritz, C., Nenzen, H.,
 231 Stanley, R. R. E., Taranu, Z. E., Gonzalez, A., Guichard, F., & Pepin, P. (2017). Signatures of the
 232 collapse and incipient recovery of an overexploited marine ecosystem. *Royal Society Open Science*, 4(7),
 233 170215. <https://doi.org/10.1098/rsos.170215>

234 Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). DeepWalk: Online learning of social representations.
 235 *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data*
 236 *Mining*, 701–710. <https://doi.org/10.1145/2623330.2623732>

237 Poisot, T., Cirtwill, A. R., Cazelles, K., Gravel, D., Fortin, M.-J., & Stouffer, D. B. (2016). The structure of
 238 probabilistic networks. *Methods in Ecology and Evolution*, 7(3), 303–312.
 239 <https://doi.org/10.1111/2041-210X.12468>

240 Poisot, T., Ouellet, M.-A., Mollentze, N., Farrell, M. J., Becker, D. J., Albery, G. F., Gibb, R. J., Seifert, S. N.,
 241 & Carlson, C. J. (2021, May 31). *Imputing the mammalian virome with linear filtering and singular*
 242 *value decomposition*. <http://arxiv.org/abs/2105.14973>

243 Poisot, T., Stouffer, D. B., & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary
 244 through space and time. *Oikos*, 124(3), 243–251. <https://doi.org/10.1111/oik.01719>

245 Ramasamy, D., & Madhoo, U. (2015). Compressive spectral embedding: Sidestepping the SVD. In C.
 246 Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in neural information*

processing systems (Vol. 28). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2015/file/4f6ffe13a5d75b2d6a3923922b3922e5-Paper.pdf>

Ray, J. C., Grimm, J., & Olive, A. (2021). The biodiversity crisis in Canada: Failures and challenges of federal and sub-national strategic and legal frameworks. *FACETS*, 6, 1044–1068. <https://doi.org/10.1139/facets-2020-0075>

Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz, N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap towards predicting species interaction networks (across space and time). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1837), 20210063. <https://doi.org/10.1098/rstb.2021.0063>

Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., & Mei, Q. (2015). LINE: Large-scale Information Network Embedding. *Proceedings of the 24th International Conference on World Wide Web*, 1067–1077. <https://doi.org/10.1145/2736277.2741093>

Turak, E., Brazill-Boast, J., Cooney, T., Drielsma, M., DelaCruz, J., Dunkerley, G., Fernandez, M., Ferrier, S., Gill, M., Jones, H., Koen, T., Leys, J., McGeoch, M., Mihoub, J.-B., Scanes, P., Schmeller, D., & Williams, K. (2017). Using the essential biodiversity variables framework to measure biodiversity change at national scale. *Biological Conservation*, 213, 264–271. <https://doi.org/10.1016/j.biocon.2016.08.019>

Wang, D., Cui, P., & Zhu, W. (2016). Structural Deep Network Embedding. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1225–1234. <https://doi.org/10.1145/2939672.2939753>

Wang, S., Arroyo, J., Vogelstein, J. T., & Priebe, C. E. (2021). Joint Embedding of Graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(4), 1324–1336. <https://doi.org/10.1109/TPAMI.2019.2948619>

Wardeh, M., Baylis, M., & Blagrove, M. S. C. (2021). Predicting mammalian hosts in which novel coronaviruses can be generated. *Nature Communications*, 12(1, 1), 780. <https://doi.org/10.1038/s41467-021-21034-5>

Wood, S. A., Russell, R., Hanson, D., Williams, R. J., & Dunne, J. A. (2015). Effects of spatial scale of sampling on food web structure. *Ecology and Evolution*, 5(17), 3769–3782.

276 <https://doi.org/10.1002/ece3.1640>

277 Yan, S., Xu, D., Zhang, B., & Zhang, H.-J. (2005). Graph embedding: A general framework for
278 dimensionality reduction. *2005 IEEE Computer Society Conference on Computer Vision and Pattern*
279 *Recognition (CVPR'05)*, 2, 830–837 vol. 2. <https://doi.org/10.1109/CVPR.2005.170>

280 Young, S. J., & Scheinerman, E. R. (2007). Random Dot Product Graph Models for Social Networks. In A.
281 Bonato & F. R. K. Chung (Eds.), *Algorithms and Models for the Web-Graph* (pp. 138–149). Springer.
282 https://doi.org/10.1007/978-3-540-77004-6_11