dendroNetwork: a R-package to create networks of dendrochronological data

true

13 February 2024

Abstract

The R-package dendroNetwork aims to make network analyses of dendrochronological data accessible for researchers. dendroNetwork enables researchers to easily construct a network of dendrochronological series based on the similarity and find patterns of matching tree-ring material. The detection of patterns is assisted by the easy application of community detection and subsequent (automated) visualization using Cytoscape. This enables dendrochronologists and other researchers to apply networks analyses to understand patterns of similarity between tree-ring series.

Introduction

Dendrochronological analyses are often based on the comparison of tree-ring series with other tree-ring series. These comparisons are often visualised as tabular data. However, to analyse the complex system of relations between tree-ring curves other methods are necessary. The author has developed a method to analyse patterns of similarity between tree-ring series using network analysis (Visser 2021b; Visser and Vorst 2022). The analyses of these papers was mostly based on scripts that are shared with the papers (Visser, n.d., n.d.). For others to apply the developed method with these scripts is possible, but cumbersome and various things need to be changed. In addition, to try the network approach on dendrochronological data for the first time can be overwhelming an seemingly difficult. To make this more easy and accessible a new package for R has been designed to help users to apply this method to their data. The package dendroNetwork enables researchers to create tables with the similarity using various measures and visualize these using R (R Core Team 2022) and Cytoscape (Otasek et al. 2019; Shannon et al. 2003).

Statement of need

There are several packages available in R for dendrochronological research (Guiterman, Lynch, and Axelson 2020; Jevšenak and Levanič 2018; Campelo, García-González, and Nabais 2012; Bunn 2008; Shi et al. 2019; Maaten-Theunissen, Maaten, and Bouriaud 2015; Reynolds, Edge, and Black 2021; Altman et al. 2014; Rademacher et al., n.d.; Campelo, Mayer, and Grabner 2019; Alday et al. 2018). An overview with links to these packages can be found online (https://github.com/RonaldVisser/Dendro_R). These R-packages in dendrochronology all fill different needs, but the nice thing is that these are all interconnected in some way in depending on each other, or that they build further into different avenues. Apart from depending on dendrochronological packages, various packages from the Tidyverse (Wickham et al. 2019) are often also needed. The various relations can easily be visualised using a network, with the edges based on the dependency of two packages on each other, and and arrow indicating the direction of the dependency.

The network shows that nearly all packages depend on dplR (Bunn 2008). The newly created package dendroNetwork fits in this ecosystem of depending packages, since it depends on both dplR and the Tidyverse. In addition, it adds a whole new world by adding network analyses through igraph (Csardi and Nepusz 2006)
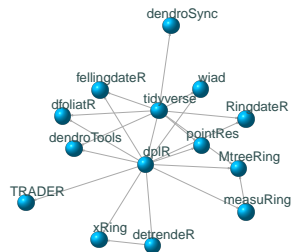
Figure 1: The interdependent relations between existing R-packages related to dendrochronology

to the ecosystem of dendrochronological packages. The igraph library has close connections to the tidyverse, creating a full circle and filling a gap as shown in the networks below.

Figure 2: The interdependent relations between R-packages including dendroNetwork. On the left only the dependencies for dendroNetwork, on the right all dendrochronological R-packages, including dendroNetwork

Workflow using dendroNetwork

The typical workflow when using the package consists of a number of successive steps. The result of each step is input for the next step.

This first step is to load the package and subsequently the dendrochronological data into the R environment using dplR:

```
library(dendroNetwork)
data(hol_rom)
```

The next step would involve calculating the similarities between each tree-ring series in the dataset.
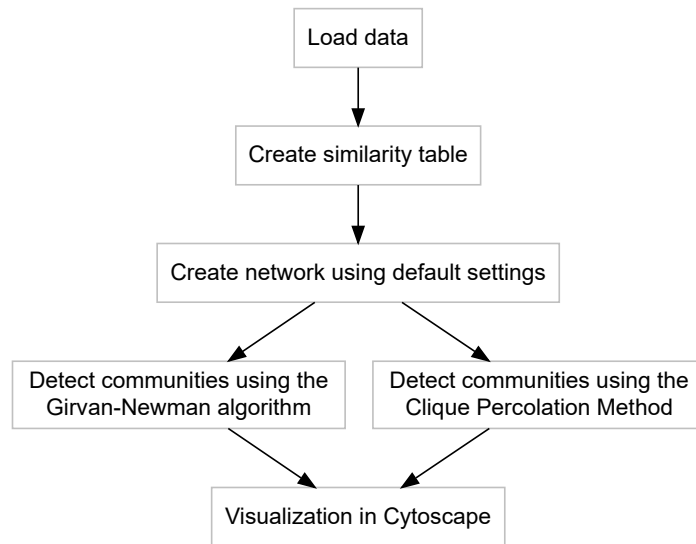
```
┌─────────────┐
│  Load data  │
└─────────────┘
       │
       ▼
┌──────────────────────┐
│ Create similarity table │
└──────────────────────┘
       │
       ▼
┌─────────────────────────────────┐
│ Create network using default settings │
└─────────────────────────────────┘
     │                    │
     ▼                    ▼
┌──────────────────────┐  ┌──────────────────────┐
│ Detect communities using the │  │ Detect communities using the │
│  Girvan-Newman algorithm     │  │  Clique Percolation Method    │
└──────────────────────┘  └──────────────────────┘
           │                    │
           ▼                    ▼
        ┌──────────────────────────┐
        │ Visualization in Cytoscape │
        └──────────────────────────┘
```

Figure 3: The typcial workflow when using dendroNetwork to visualise a network

```r
sim_table_hol <- sim_table(hol_rom)
```

In the next step edges of the network are created based on the similarity. Each similarity between two curves above certain threshold settings will result in an edge in the network. The default settings are 0.5 for the correlation, 0.7 for the Synchronous Growth Changes or sgc (Visser 2021a) with a probability of exceedence below 0.0001.

```r
g_hol <- dendro_network(sim_table_hol)
```

The next steps will be to find communities using either the Girvan-Newman algorithm (Girvan and Newman 2002) or clique percolation method (Palla et al. 2005) , or both.

```r
g_hol_gn <- gn_names(g_hol)
g_hol_cpm <- clique_community_names(g_hol, k=3)
hol_com_cpm_all <- find_all_cpm_com(g_hol)
```

The next step is to visualize and explore the networks using Cytoscape or using R. Various functions are available to create visual styles for the communities in Cytoscape from R.

```r
plot(g_hol, vertex.color="deepskyblue", vertex.size=15, vertex.frame.color="gray",
     vertex.label.color="darkslategrey", vertex.label.cex=0.8, vertex.label.dist=2)
```
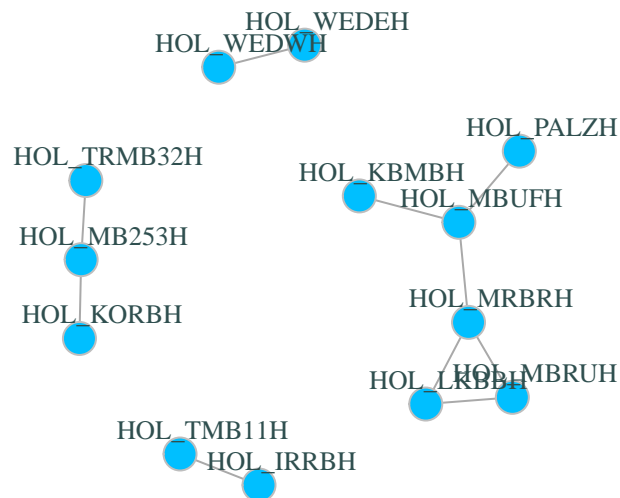

Figure 4: Network of the dendrochronological site chronologies in hol_rom

The main advantage is that visualisation in Cytoscape is more easy, intuitive and visual. In addition, it is very easy to automate workflows in Cytoscape with R (using RCy3). For this purpose we need to start Cytoscape firstly. After Cytoscape has completely loaded, the next steps can be taken.

1. The network can now be loaded in Cytoscape for further visualisation: `cyto_create_graph(g_hol, CPM_table = hol_com_cpm_all, GN_table = g_hol_gn)`
2. Styles for visualisation can now be generated. However, Cytoscape comes with a lot of default styles that can be confusing. Therefore it is recommended to use: `cyto_clean_styles()` once in a session.
3. To visualize the styles for CPM with only k=3: `cyto_create_cpm_style(g_hol, k=3, com_k = g_hol_cpm)`

   - This can be repeated for all possible clique sizes. To find the maximum clique size in a network, please use: `igraph::clique_num(g_hol)`.
   - To automate this: `for (i in 3:igraph::clique_num(g_hol)) { cyto_create_cpm_style(g_hol, k=i, com_k = g_hol_cpm)}`.

4. To visualize the styles using the Girvan-Newman algorithm (GN): `cyto_create_gn_style(g_hol)` This would look something like this in Cytoscape:
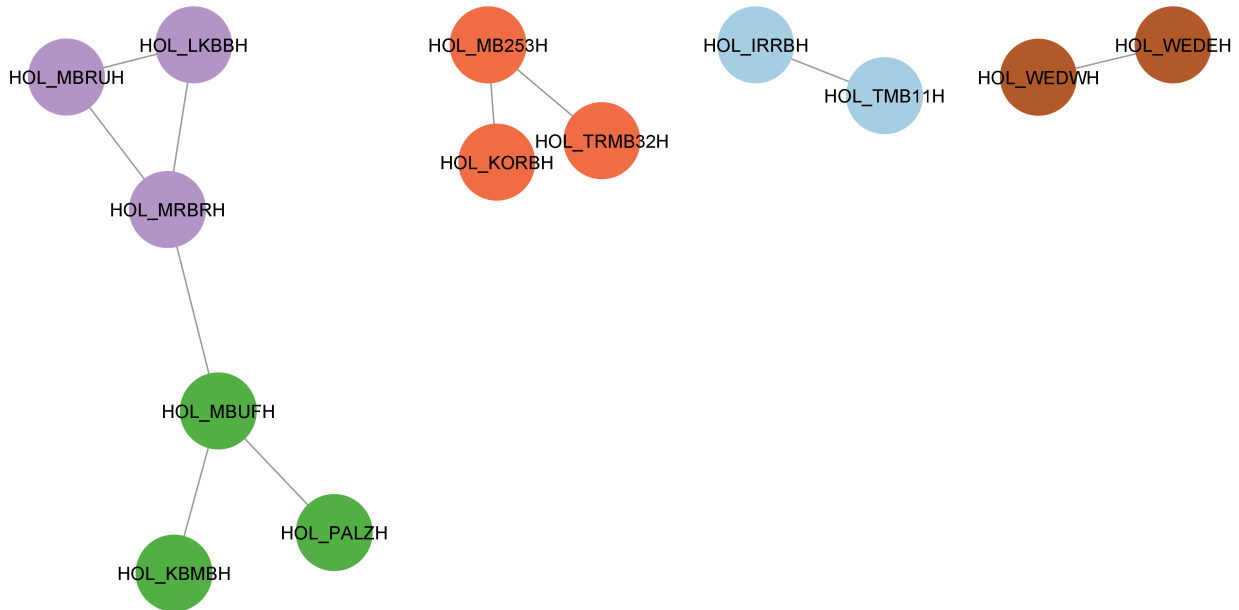

Figure 5: The network of Roman sitechronologies with the Girvan-Newman communities visualized using Cytoscape.

Conclusion

The new R package dendroNetwork has been developed based on earlier research by the author. To enable other researchers to use network analyses on dendrochronological data this new package fills a gap that

exists in the current network of R-packages related to dendrochronology by also connecting igraph to the existing packages in the discipline. The creation of networks in R is made easy with various functions and the visualization in Cytoscape is automated in R, enabling the researcher to quickly visualise and analyse the resulting networks in an intuitive manner. The package is easy to use and enables reproducible network analyses within dendrochronology.

Acknowledgements

References

Alday, Josu G., Tatiana A. Shestakova, Víctor Resco de Dios, and Jordi Voltas. 2018. "DendroSync: An r Package to Unravel Synchrony Patterns in Tree-Ring Networks." *Dendrochronologia* 47 (February): 17–22. https://doi.org/10.1016/j.dendro.2017.12.003.

Altman, J., P. Fibich, J. Dolezal, and T. Aakala. 2014. "TRADER: A Package for Tree Ring Analysis of Disturbance Events in r." *Dendrochronologia* 32 (2): 107–12. https://doi.org/10.1016/j.dendro.2014.01.004.

Bunn, Andrew G. 2008. "A Dendrochronology Program Library in r (dplR)." *Dendrochronologia* 26 (2): 115124. https://doi.org/https://doi.org/10.1016/j.dendro.2008.01.002.

Campelo, Filipe, Ignacio García-González, and Cristina Nabais. 2012. "detrendeR - a Graphical User Interface to Process and Visualize Tree-Ring Data Using r." *Dendrochronologia* 30: 57–60. https://doi.org/16/j.dendro.2011.01.010.

Campelo, Filipe, Konrad Mayer, and Michael Grabner. 2019. "xRing—An R Package to Identify and Measure Tree-Ring Features Using X-Ray Microdensity Profiles." *Dendrochronologia* 53 (February): 17–21. https://doi.org/10.1016/j.dendro.2018.11.002.

Csardi, Gabor, and Tamas Nepusz. 2006. "The Igraph Software Package for Complex Network Research." *InterJournal Complex Systems* 1695 (5): 1–9. https://igraph.org.

Girvan, M., and M. E. J. Newman. 2002. "Community Structure in Social and Biological Networks." *Proceedings of the National Academy of Sciences of the United States of America* 99 (12): 7821–26. https://doi.org/10.1073/pnas.122653799.

Guiterman, Christopher H., Ann M. Lynch, and Jodi N. Axelson. 2020. "dfoliatR: An R Package for Detection and Analysis of Insect Defoliation Signals in Tree Rings." *Dendrochronologia* 63 (October): 125750. https://doi.org/10.1016/j.dendro.2020.125750.

Jevšenak, Jernej, and Tom Levanič. 2018. "dendroTools: R Package for Studying Linear and Nonlinear Responses Between Tree-Rings and Daily Environmental Data." *Dendrochronologia* 48 (April): 32–39. https://doi.org/10.1016/j.dendro.2018.01.005.

Maaten-Theunissen, Marieke van der, Ernst van der Maaten, and Olivier Bouriaud. 2015. "pointRes: An r Package to Analyze Pointer Years and Components of Resilience." *Dendrochronologia* 35 (October): 34–38. https://doi.org/10.1016/j.dendro.2015.05.006.

Otasek, David, John H. Morris, Jorge Bouças, Alexander R. Pico, and Barry Demchak. 2019. "Cytoscape Automation: Empowering Workflow-Based Network Analysis." *Genome Biology* 20 (1): 185. https://doi.org/10.1186/s13059-019-1758-4.

Palla, Gergely, Imre Derenyi, Illes Farkas, and Tamas Vicsek. 2005. "Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society." *Nature* 435 (7043): 814–18. https://doi.org/10.1038/nature03607.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. https://www.r-project.org.

Rademacher, Tim, Bijan Seyednasrollah, David Basler, Jian Cheng, Tessa Mandra, Elise Miller, Zuid Lin, et al. n.d. "The Wood Image Analysis and Dataset (WIAD): Open-Access Visual Analysis Tools to Advance the Ecological Data Revolution." *Methods in Ecology and Evolution* n/a (n/a). https://doi.org/10.1111/2041-210X.13717.

Reynolds, David J., David C. Edge, and Bryan A. Black. 2021. "RingdateR: A Statistical and Graphical Tool for Crossdating." *Dendrochronologia* 65 (February): 125797. https://doi.org/10.1016/j.dendro.2020.125797.

Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." *Genome Research* 13 (11): 2498–2504. https://doi.org/10.1101/gr.1239303.

Shi, Jingning, Wei Xiang, Qijing Liu, and Sher Shah. 2019. "MtreeRing: An R Package with Graphical User Interface for Automatic Measurement of Tree Ring Widths Using Image Processing Techniques." *Dendrochronologia* 58 (December): 125644. https://doi.org/10.1016/j.dendro.2019.125644.

Visser, Ronald M. 2021a. "On the Similarity of Tree-Ring Patterns: Assessing the Influence of Semi-Synchronous Growth Changes on the Gleichläufigkeitskoeffizient for Big Tree-Ring Data Sets." *Archaeometry* 63 (1): 204–15. https://doi.org/https://doi.org/10.1111/arcm.12600.

———. 2021b. "Dendrochronological Provenance Patterns. Network Analysis of Tree-Ring Material Reveals Spatial and Economic Relations of Roman Timber in the Continental North-Western Provinces." *Journal of Computer Applications in Archaeology* 4 (1): 230253. https://doi.org/10.5334/jcaa.79.

———. n.d. "Dendrochronological Provenance Patterns. Code and Data of Network Analysis of Tree-Ring Material." https://doi.org/10.5281/zenodo.10200361.

Visser, Ronald M., and Yardeni Vorst. 2022. "Connecting Ships: Using Dendrochronological Network Analysis to Determine the Wood Provenance of Roman-Period River Barges Found in the Lower Rhine Region and Visualise Wood Use Patterns." *International Journal of Wood Culture* 3 (1-3): 123–51. https://doi.org/10.1163/27723194-bja10014.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.