

Minimising Crossings in a Tree-Based Network

Jonathan Klawitter¹[0000-0001-8917-5269] and Peter
Stumpf²[0000-0003-0531-9769]

¹ School of Computer Science, University of Auckland, New Zealand,
jo.klawitter@gmail.com

² Faculty of Computer Science and Mathematics, University of Passau, Germany,
Peter.Stumpf@uni-passau.de

Abstract. A tree-based network N is a rooted graph that has a spanning tree T which is the subdivision of a rooted binary tree. We show that crossing minimisation for drawings of N is either NP-hard or polynomial time solvable depending on the drawing style used for non- T edges.

1 Introduction

Phylogenetic trees and *networks* are rooted, leaf-labelled graphs used to model and visualise the evolutionary history of a set of taxa, e.g. species or languages [3]. A *tree-based network* N is a phylogenetic network with a spanning tree T that is the subdivision of a binary phylogenetic tree [1]. For fixed T , we refer to edges not covered by T in N as *cross edges*. We assume that vertices of N are assigned a height such that leaves have height zero and otherwise only two endpoints of a cross edge may have the same height. We only consider drawings of N where T is planar. Motivated by examples in the literature [4, 5] we identified the following drawing styles for cross edges. First, cross edges may be drawn x -monotone such that they are *horizontal*, *snakes* (i.e. have to bends), *curves*, or *straight lines*. Second, we consider the *ear* and *ear** drawing styles where a cross edge (u, v) is drawn with two bends such that the vertical segment is to the right of the subtree containing u and v or of T , respectively. See Figure 1 (a)–(d) for examples.

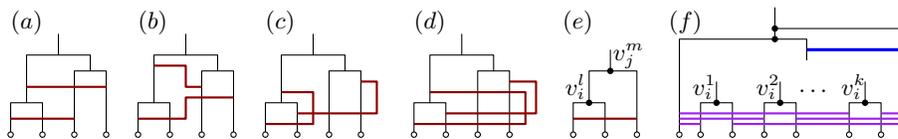


Fig. 1. (a)–(d): The horizontal, snake, ear, and ear* drawing styles, respectively; (e), (f): edge and vertex gadgets.

2 Horizontal drawing style

Theorem 1. *Crossing minimisation for a tree-based network with the horizontal drawing style is NP-complete.*

We sketch a polynomial-time reduction of an instance G of MAX-CUT to an instance of crossing minimisation of a tree-based network N . Assume without loss of generality that G contains no vertex of degree 1. Sort $V(G)$ in an arbitrary order, say v_1, v_2, \dots, v_n .

We build the base tree T of N as follows. A gadget for each edge $\{v_i, v_j\}$, $i < j$, as shown in Figure 1 (e), where v_i^l (v_j^m) is a representative of v_i (resp. v_j), induces zero crossings if v_i, v_j are rotated differently and one crossing otherwise. Thus it represents whether the edge is in the cut or not. We connect the edge gadgets with a binary tree. For $i \in \{1, \dots, n\}$, let v_i^1, \dots, v_i^k be all vertices corresponding to $v_i \in V(G)$. We give v_i^l the height $i + l\epsilon$ for sufficiently small ϵ . Furthermore, to force them to all have the “same” rotation we add two vertices above the so far build tree and bundles of horizontal edges as shown in Figure 1 (f). Further bundles of horizontal edges fix the part of T that connects the edge gadgets and the upper part of vertex gadgets.

Note that several of the drawing styles can be reduced to the horizontal drawing style, which gives us the following corollary.

Corollary 1. *Crossing minimisation for a tree-based network with the snake, monotone curve, or straight line drawing style is NP-complete.*

3 Ears drawing style

Theorem 2. *Crossing minimisation for a tree-based network N on n vertices and k cross edges drawn with the ears drawing style can be solved in $\mathcal{O}(nk)$ time.*

Let v be a vertex of T . The key observation is that the rotation of v only determines crossings of edges from its left subtree through its right subtree or vice versa. Thus the best rotation can be determined independently for each vertex of T . The following algorithm runs in $\mathcal{O}(nk)$ time. First we determine the lowest common ancestor (lca) for each pair of endpoints of a cross edge in $\mathcal{O}(n)$ time [2]. Then we sweep from the leaves of T towards the root. At every endpoint v of a cross edge we determine for every vertex u of T the width of its left and right subtree at the height of v in $\mathcal{O}(n)$ time. Then from v up to its cross edge lca, we add up for each vertex of T the width of the subtree not containing v to a counter. When reaching a vertex u of T , we can decide its best rotation based on this counter. Lastly, we extend a partial order of groups of nested vertical segments to a total order to minimise crossings between cross edges.

Adjusting this algorithm to propagate the height up to the root instead of only to the lca, we get the following corollary.

Corollary 2. *Crossing minimisation for a tree-based network N on n vertices and k cross edges drawn with the ears* drawing style can be solved in $\mathcal{O}(nk)$ time.*

Acknowledgements

The first author thanks the New Zealand Marsden Fund for their financial support. We thank the anonymous referees.

References

1. Francis, A.R., Steel, M.: Which Phylogenetic Networks are Merely Trees with Additional Arcs? *Systematic Biology* **64**(5), 768–777 (2015). <https://doi.org/10.1093/sysbio/syv037>
2. Gabow, H.N., Tarjan, R.E.: A linear-time algorithm for a special case of disjoint set union. *Journal of Computer and System Sciences* **30**(2), 209–221 (1985). [https://doi.org/10.1016/0022-0000\(85\)90014-5](https://doi.org/10.1016/0022-0000(85)90014-5)
3. Huson, D.H., Rupp, R., Scornavacca, C.: *Phylogenetic networks: concepts, algorithms and applications*. Cambridge University Press (2010). <https://doi.org/10.1093/sysbio/syr055>
4. Kumar, V., Lammers, F., Bidon, T., Pfenninger, M., Kolter, L., Nilsson, M.A., Janke, A.: The evolutionary history of bears is characterized by gene flow across species. *Scientific Reports* **7**(1) (2017). <https://doi.org/10.1038/srep46487>
5. Vaughan, T.G., Welch, D., Drummond, A.J., Biggs, P.J., George, T., French, N.P.: Inferring Ancestral Recombination Graphs from Bacterial Genomic Data. *Genetics* **205**(2), 857–870 (2017). <https://doi.org/10.1534/genetics.116.193425>