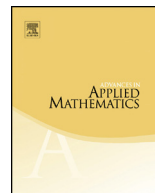




ELSEVIER

Contents lists available at ScienceDirect

Advances in Applied Mathematics

[www.elsevier.com/locate/yaama](http://www.elsevier.com/locate/yaama)

## New characterisations of tree-based networks and proximity measures

Andrew Francis<sup>a</sup>, Charles Semple<sup>b</sup>, Mike Steel<sup>b,\*</sup><sup>a</sup> *Centre for Research in Mathematics, School of Computing, Engineering and Mathematics, Western Sydney University, Sydney, Australia*<sup>b</sup> *School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand*

### ARTICLE INFO

#### Article history:

Received 14 November 2016

Received in revised form 22 August 2017

Accepted 23 August 2017

Available online xxxx

#### MSC:

05C05

05C20

05C85

05D15

92D15

#### Keywords:

Phylogenetic network

Tree-based network

Antichain

Path partition

Dilworth's theorem

### ABSTRACT

Phylogenetic networks are a type of directed acyclic graph that represent how a set  $X$  of present-day species are descended from a common ancestor by processes of speciation and reticulate evolution. In the absence of reticulate evolution, such networks are simply phylogenetic (evolutionary) trees. Moreover, phylogenetic networks that are not trees can sometimes be represented as phylogenetic trees with additional directed edges placed between their edges. Such networks are called *tree-based*, and the class of phylogenetic networks that are tree-based has recently been characterised. In this paper, we establish a number of new characterisations of tree-based networks in terms of path partitions and antichains (in the spirit of Dilworth's theorem), as well as via matchings in a bipartite graph. We also show that a temporal network is tree-based if and only if it satisfies an antichain-to-leaf condition. In the second part of the paper, we define three indices that measure the extent to which an arbitrary phylogenetic network deviates from being tree-based. We describe how

\* Corresponding author.

*E-mail addresses:* [a.francis@westernsydney.edu.au](mailto:a.francis@westernsydney.edu.au) (A. Francis), [charles.semple@canterbury.ac.nz](mailto:charles.semple@canterbury.ac.nz) (C. Semple), [mike.steel@canterbury.ac.nz](mailto:mike.steel@canterbury.ac.nz) (M. Steel).

<http://dx.doi.org/10.1016/j.aam.2017.08.003>

0196-8858/© 2017 Elsevier Inc. All rights reserved.

these three indices can be computed efficiently using classical results concerning maximum-sized matchings in bipartite graphs.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

Phylogenetic networks are of increasing interest in the literature as they allow for the representation of reticulate (non-tree-like) processes in evolution. From a mathematical perspective, there are several particularly attractive classes of phylogenetic networks. One of those classes is tree-based networks. Intuitively, a phylogenetic network is tree-based if it can be obtained from a phylogenetic tree  $\mathcal{T}$  by simply adding edges whose end-vertices subdivide edges of  $\mathcal{T}$ . Formalised and studied in [5], tree-based networks have since been studied in a number of recent papers [1,7,9,11,12], in a variety of contexts.

In this paper, we establish several new characterisations of tree-based networks. These characterisations are based on antichains, path partitions, and matchings in bipartite graphs, and complement the previous characterisations based on bipartite matchings [9, 12]. Furthermore, with the aid of these characterisations, we explore indices quantifying the closeness of an arbitrary phylogenetic network  $\mathcal{N}$  to being tree-based. Each of the considered indices is computable in time polynomial in the size of  $\mathcal{N}$ , that is, in the number of vertices in  $\mathcal{N}$ , by finding a maximum-size matching in certain bipartite graphs based on  $\mathcal{N}$ .

The paper is organised as follows. The rest of the introduction contains some formal definitions and previous results. In Section 2, we state the new characterisations of tree-based networks as well as a characterisation of tree-based networks within the class of temporal networks. Except for one characterisation which is proved in Section 4, the proofs of these characterisations are given in Section 3. In Section 4, we consider three indices quantifying the extent to which an arbitrary phylogenetic network is tree-based. We end the paper with a brief conclusion and further questions in Section 5.

### 1.1. Definitions

Throughout the paper,  $X$  denotes a non-empty finite set. A *phylogenetic network*  $\mathcal{N}$  on  $X$  is a rooted acyclic digraph with no edges in parallel and satisfying the following properties:

- (i) the (unique) root has out-degree two;
- (ii) a vertex with out-degree zero has in-degree one, and the set of vertices with out-degree zero is  $X$ ; and
- (iii) all other vertices either have in-degree one and out-degree two, or in-degree two and out-degree one.

Download English Version:

<https://daneshyari.com/en/article/5775352>

Download Persian Version:

<https://daneshyari.com/article/5775352>

[Daneshyari.com](https://daneshyari.com)