Published in *Journal of classification*, 2007, vol. 24, no. 2, p.235-249
which should be cited to refer to this work

Journal of Classification manuscript No.
(will be inserted by the editor)

The Metric Bridge Partition Problem

Partitioning of a metric space into two subspaces linked by an edge in any optimal realization

Alain Hertz · Sacha Varone

Received: date / Accepted: date

Abstract Let $G = (V, E, w)$ be a graph with vertex and edge sets V and E, respectively, and $w : E \rightarrow I\!R^+$ a function which assigns a positive weigth or length to each edge of G. G is called a realization of a finite metric space (M, d), with $M = \{1, ..., n\}$ if and only if $\{1, ..., n\} \subseteq V$ and $d(i, j)$ is equal to the length of the shortest chain linking i and j in G $\forall i, j = 1, ..., n$. A realization G of (M, d), is said optimal if the sum of its weights is minimal among all the realizations of (M, d). Consider a partition of M into two nonempty subsets K and L, and let e be an edge in a realization G of (M, d); we say that e is a bridge linking K with L if e belongs to all chains in G linking a vertex of K with a vertex of L. The Metric Bridge Partition Problem is to determine if the elements of a finite metric space (M, d) can be partitioned into two nonempty subsets K and L such that all optimal realizations of (M, d) contain a bridge linking K with L. We prove in this paper that this problem is polynomially solvable. We also describe an algorithm that constructs an optimal realization of (M, d) from optimal realizations of $(K, d|_K)$ and $(L, d|_L)$.

Keywords metric, partition, optimal realization, decomposition, algorithm

1 Introduction

A *metric space* is a couple (M, d) such that M is a set and d is a metric, which is a positive and reflexive function, with respect to the triangular inequalities (i.e. d is a function defined on $M \times M$ such that $d(x, y) = d(y, x) > 0$ $\forall x \neq y$, $d(x, x) = 0$ $\forall x$, and $d(x, z) \leq d(x, y) + d(y, z)$ $\forall x, y, z$). Moreover, (M, d) is a finite metric space if M has a finite number of elements.

This work has been supported by grant PA002-104974/1 from the Swiss National Science Foundation, received by S. Varone.

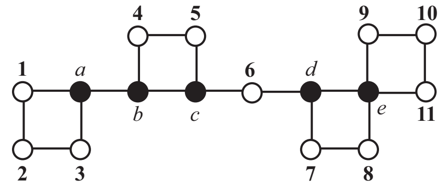
A. Hertz
Département de mathématiques et de génie industriel, École Polytechnique, CP 6079, succ. Centre-ville, Montréal (QC), H3C 3A7, Canada, E-mail: alain.hertz@gerad.ca

S. Varone
Haute École de Gestion, Campus Battelle de la HEG, Bâtiment F, Route de Drize 7, 1227 Carouge, Switzerland, Tel:+41 22 388 1700. E-mail: sacha.varone@hesge.ch

Let $G = (V, E, w)$ be a graph, with vertex and edge sets V and E, respectively, and $w : E \to I\!R^+$ a function which assigns a positive weight or length to each edge of G. Furthermore, let $d^G(i, j)$ denote the length of a shortest chain in G linking vertices i and j. We say that G is a realization of a finite metric space (M, d), with $M = \{1, ..., n\}$ if and only if $\{1, ..., n\} \subseteq V$ and $d^G(i, j) = d(i, j)\ \forall i, j = 1, ..., n$. The vertices which do not belong to M are called *auxiliary vertices*. A realization of (M, d) is called optimal when the sum of its weights is minimal among all the realizations of (M, d). For illustration, a metric space together with an optimal realization G are shown in Figure 1. All edges of the graph have length one, and the black points a, b, c, d, e are five auxiliary vertices while the white ones are the elements of M.

$$
\begin{array}{c|ccccccccccc}
 & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} & \mathbf{8} & \mathbf{9} & \mathbf{10} & \mathbf{11} \\
\hline
\mathbf{1} & 0 & 1 & 2 & 3 & 4 & 4 & 6 & 7 & 7 & 8 & 7 \\
\mathbf{2} & & 0 & 1 & 4 & 5 & 5 & 7 & 8 & 8 & 9 & 8 \\
\mathbf{3} & & & 0 & 3 & 4 & 4 & 6 & 7 & 7 & 8 & 7 \\
\mathbf{4} & & & & 0 & 1 & 3 & 5 & 6 & 6 & 7 & 6 \\
\mathbf{5} & & & & & 0 & 2 & 4 & 5 & 5 & 6 & 5 \\
\mathbf{6} & & & & & & 0 & 2 & 3 & 3 & 4 & 3 \\
\mathbf{7} & & & & & & & 0 & 1 & 3 & 4 & 3 \\
\mathbf{8} & & & & & & & & 0 & 2 & 3 & 2 \\
\mathbf{9} & & & & & & & & & 0 & 1 & 2 \\
\mathbf{10} & & & & & & & & & & 0 & 1 \\
\mathbf{11} & & & & & & & & & & & 0
\end{array}
$$

Upper diagonal of the distance matrix of a metric space (M,d) with $M=\{1,...,11)$



Optimal realization of (M,d)

Figure 1. A metric space with an optimal realization

The embedding of finite metric spaces in graphs has applications in varied fields as computational biology (Landry, Lapointe and Kirsch, 1996; Makarenkov, 2002) (e.g., constructing phylogenetic trees from genetic distances among living species), electrical networks (Hakimi and Yau, 1964), coding techniques (Dress, 1984), psychology (Cunningham, 1978), internet tomography (Chung, Garret, Graham and Shallcross, 2001), and compression softwares (Li, Chen, Ma and Vitányi, 2004).

The problem of finding optimal realizations of metric spaces was first proposed by Hakimi and Yau (1964) who also gave a polynomial algorithm for the special case where the metric space has a realization as a tree. While every finite metric space has an optimal realization (Imrich and Stockiĭ, 1972; Imrich, Simões-Pereira and Zamfirescu, 1984), finding such realizations is an NP-hard problem (Winkler, 1988).

Optimal realizations can be constructed using building blocks. More precisely, for a graph G, we recall that a *cutpoint*, respectively a *bridge*, is a vertex, respectively an *edge*, whose removal strictly increases the number of connected component of G; a *block* is a maximal two-connected subgraph or a bridge in G. Imrich, Simões-Pereira and Zamfirescu (1984) have proved the following theorem.

Theorem 1 *(Imrich, Simões-Pereira and Zamfirescu, 1984) Let G be a realization of a finite metric space (M, d), let $G_1, \cdots, G_k$ be the blocks of G, and let M_i be the union of the points of M in G_i together with the cutpoints of G in G_i. If every G_i is an optimal realization of the metric space induced by G on M_i, then G is also optimal.*

For example an optimal realization of the metric space of Figure 1 can be obtained by putting together optimal realizations of the metric spaces induced on $\{1, 2, 3, a\}$, $\{a, b\}$, $\{4, 5, b, c\}$, $\{6, c\}$, $\{6, d\}$, $\{7, 8, d, e\}$, and $\{9, 10, 11, e\}$.

It is therefore interesting to be able to recognize metric spaces which contain at least one bridge in all optimal realizations. This is exactly the topic of our paper. More precisely, consider a partition of M into two nonempty subsets K and L, and let e be an edge in a realization G of (M, d). We say that e is a bridge linking K with L if e belongs to all chains in G linking a vertex of K with a vertex of L. The *Metric Bridge Partition Problem* is to determine if the elements of a given finite metric space (M, d) can be partitioned into two nonempty subsets K and L such that all optimal realizations of (M, d) contain a bridge linking K with L. For example, on the basis of the distance matrix of Figure 1 (and without any knowledge of the optimal realization), we would like to be able to state that all optimal realizations contain a bridge linking $K = \{1, 2, 3, 4, 5, 6\}$ with $L = \{7, 8, 9, 10, 11\}$, or $K = \{1, 2, 3, 4, 5\}$ with $L = \{6, 7, 8, 9, 10, 11\}$, or $K = \{1, 2, 3\}$ with $L = \{4, 5, 6, 7, 8, 9, 10, 11\}$. We prove in this paper that the Metric Bridge Partition Problem is polynomially solvable.

2 Definitions and Known Results

It is well-known that the unique optimal realization of a metric space on three points i, j, k is a tree T. The *hub* of i, j, k, denoted h_{ijk}, is the junction point which induces the shortest path between i, j and k. It is defined as follows:

$$\begin{aligned} d^T(h_{ijk}, i) &= \tfrac{1}{2}(d(i,j) + d(i,k) - d(j,k)), \\ d^T(h_{ijk}, j) &= \tfrac{1}{2}(d(j,i) + d(j,k) - d(i,k)), \\ d^T(h_{ijk}, k) &= \tfrac{1}{2}(d(k,i) + d(k,j) - d(i,j)). \end{aligned}$$

Assume that the distance $d(i, j)$ is larger than or equal to $d(i, k)$ and $d(j, k)$. If $d(i, j) < d(i, k) + d(j, k)$, then T has three leaves i, j and k, and one auxiliary vertex corresponding to the hub h_{ijk}, else T is a chain linking i and j that traverses $k = h_{ijk}$ (see Figure 2).

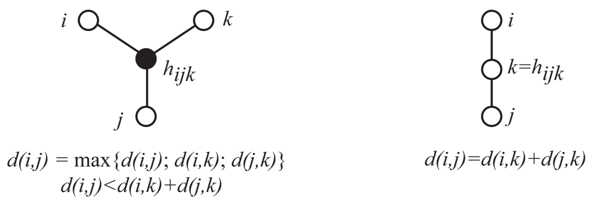


Figure 2. Optimal realizations of three points

Let $s_{ijk\ell}$ denote the sum $d(i, j) + d(k, \ell)$. It is also well-known that the optimal realization of a metric space on four points i, j, k, ℓ is a unique tree if and only if two of the sums $s_{ijk\ell}, s_{ikj\ell}, s_{i\ell jk}$ are equal and not smaller than the third. Moreover, if $s_{ijk\ell} < s_{ikj\ell} = s_{i\ell jk}$, then the tree has a bridge $(h_{ijk}, h_{ik\ell})$ of length $s_{ikj\ell} - s_{ijk\ell} > 0$ linking $\{i, j\}$ with $\{k, \ell\}$. The three possible configurations are represented in Figure 3 (the other cases are equivalent).

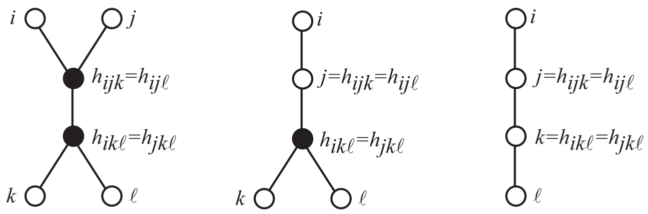


Figure 3. Optimal realizations of four points

Definition 1 A finite metric space (M, d) is *reducible* if and only if all its optimal realizations contain a vertex of degree one (i.e, a vertex with exactly one neighbor).

In other words (see for example Imrich, Simões-Pereira and Zamfirescu (1984)), a finite metric space (M, d) is reducible if and only if M contains an element i, called *endpoint*, such that $d(i, j) + d(i, k) - d(j, k) > 0$ for all $j, k \neq i$. An optimal realization of a reducible metric space (M, d) can easily be obtained from an optimal realization of a metric space (M', d') which has fewer endpoints or fewer elements than M. More precisely, consider an endpoint i in a reducible metric space (M, d), and define $\alpha = \min\{\frac{1}{2}(d(i, j) + d(i, k) - d(j, k))\}$, the minimum being taken over all $j, k \neq i$. There are two possible cases:

- If there is an element $j \in M$ with $d(i, j) = \alpha$, then set M' equal to $M \backslash \{i\}$, and set $d' = d|_{M'}$ (i.e., d' is the distance matrix induced by d on M'). An optimal realization of (M, d) can be obtained from an optimal realization of (M', d') by adding a vertex i and an edge of length α linking i with j.
- If there is no element $j \in M$ with $d(i, j) = \alpha$, then set $M' = M \backslash \{i\} \cup \{a\}$ and define $d'(j, k) = d(j, k)$ for all $j, k \neq a$, $d'(a, j) = d(i, j) - \alpha$ for all $j \neq a$. An optimal realization of (M, d) can be obtained from an optimal realization of (M', d'), by adding a vertex i and an edge of length α linking i with a.

Definition 2 Consider a finite metric space (M, d), a partition of M into two non-empty subsets K, L and a mapping $f : M \to I\!R^+$. The triplet (K, L, f) is said *nice* if

- $d(x, y) \leq f(x) + f(y)$ for all x, y in M, equality holding whenever $x \in K$ and $y \in L$, and
- $f(x) > 0$ at least once in K and once in L.

The above definition is motivated by the following result proved by Imrich and Stockiĭ (1972) and Imrich, Simões-Pereira and Zamfirescu (1984)

Theorem 2 *(Imrich and Stockiĭ, 1972; Imrich, Simões-Pereira and Zamfirescu, 1984) Suppose (M, d) is a finite metric space to which there exists a nice triplet (K, L, f). Then every optimal realization G of (M, d) has a cut-point c or a bridge with a point c on it such that all chains linking K with L go through c, and $d^G(x, c) = f(x) \quad \forall x \in M$.*

3 New Results and algorithms

We give two results, which will justify our decomposition algorithm 1. Only a sketch of the proofs is given here; a complete proof of each theorem is available in the Appendix.

The following theorem provides a sufficient condition for the existence of a bridge in every optimal realization of a finite metric space (M, d). It states that if a partition of the metric space into two parts satisfies some properties, then there is a bridge between those two parts in every optimal realization.

Theorem 3 *Suppose (M, d) is a finite metric space to which there exists a partition of M into two non-empty subsets K, L with $|K| > 1$ and $|L| > 1$, and assume*
(1) $s_{ijk\ell} < s_{ikj\ell} = s_{i\ell jk} \quad \forall i, j \in K$ *and* $k, \ell \in L$.
(2) $\exists\, x, y \in K$ *and* $z, t \in L$ *such that*
$s_{xzyt} - s_{xyzt} \leq s_{ikj\ell} - s_{ijk\ell} \quad \forall i, j \in K$ *and* $k, \ell \in L$
Then every optimal realization of (M, d) has a bridge (h_{xyz}, h_{xzt}) linking K with L.

In order to prove this result, we build two distance functions f and g that measure the distance from the borders of the bridge (if any) to a vertex in K or in L. We show that (K, L, f) and (K, L, g) are both nice triplets, which implies the existence of a bridge in every optimal realization.

A necessary condition for the existence of a bridge is given by the next theorem, whose proof is rather technical (see Appendix for the proof).

Theorem 4 *Suppose (M, d) is an irreducible finite metric space. If there is a partition of M into two non-empty subsets K, L such that all optimal realizations of (M, d) contain a bridge (u, v) linking K with L, then*
(1) $|K| > 1$ *and* $|L| > 1$,
(2) $s_{ijk\ell} < s_{ikj\ell} = s_{i\ell jk} \quad \forall i, j \in K$ *and* $k, \ell \in L$,
(3) $\exists\, x, y \in K$ *and* $z, t \in L$ *such that*
•$s_{xzyt} - s_{xyzt} \leq s_{ikj\ell} - s_{ijk\ell} \quad \forall i, j \in K$ *and* $k, \ell \in L$
•$i \in K \Leftrightarrow d(x, i) - d(z, i) \leq d(x, y) - d(z, y)$.

Algorithm $MetricBridgePartition$ determines if a given finite irreducible metric space (M, d) contains a bridge.

Theorem 5 *The $MetricBridgePartition$ algorithm works correctly and is polynomial.*

Proof Correctness of the algorithm follows from the results of the two previous theorems. Indeed, if the algorithm stops with four elements x, y, z, t and a partition of M into two sets K and $M \backslash K$, then properties (1) and (2) of Theorem 3 are satisfied, and we conclude that every optimal realization of (M, d) has a bridge linking K with L. Moreover, if there exists a partition of M into two sets K and L such that every optimal realization of (M, d) has a bridge, then we know from Theorem 4 that such a partition will be found.
The first loop can be simplified by choosing x (or any variable among x, y, z, t) at the beginning of the algorithm, and has therefore a complexity of $O(|M|^3)$. The inner loop is done in $O(|M|)$. The condition on the sums can be checked in $O(|M|^3)$ by

Algorithm 1 *MetricBridgePartition*

Require: A finite irreducible metric space (M, d);
Ensure: Four elements $x, y, z, t \in M$ and a set K such that there is a bridge linking K with $M \backslash K$, or a message indicating that no optimal realization of (M, d) has a bridge;

```
for all x, y, z, t ∈ M such that s_xyzt < s_xzyt = s_xtyz do
   K ← {x, y} and L ← {z, t};
   for all i ∈ M\{x, y, z, t} do
      if d(x, i) − d(z, i) ≤ d(x, y) − d(z, y) then
         K ← K ∪ {i}
      else
         L ← L ∪ {i}
      end if
   end for
   if s_xzyt − s_xyzt ≤ s_ikjℓ − s_ijkℓ and s_ijkℓ < s_ikjℓ = s_iℓjk ∀i, j ∈ K k, ℓ ∈ L then
      STOP: return x, y, z, t and K.
   end if
end for
return a message indicating that no optimal realization of (M, d) has a bridge.
```

choosing i (or any variable among i, j, k, l). All other instructions in the algorithm are done in constant time. Therefore, the algorithm is polynomial since its complexity is $O(|M|^6)$ $\square$.

The $MetricBridgePartition$ algorithm can be used to decompose a given finite metric space (M, d) into metric spaces $(M_1, d_1), \cdots, (M_r, d_r)$ such that no optimal realization of (M_i, d_i) $(i = 1, \cdots, r)$ has a bridge. According to Theorem 1, an optimal realization of (M, d) can then be obtained by connecting optimal realizations of $(M_1, d_1), \cdots, (M_r, d_r)$ with bridges. More precisely, assume the existence of the three following algorithms:

- algorithm $NoBridge$ constructs an optimal realization of a finite metric space if such a realization has no bridge;
- algorithm *Reduce* transforms any finite reducible metric space (M, d) into an irreducible metric space (M', d');
- given a finite reducible metric space (M, d) and an irreducible metric space (M', d') obtained by applying *Reduce* on (M, d), and given also an optimal realization G' of (M', d'), algorithm *Extend* constructs an optimal realization G of (M, d).

As explained in Section 2, algorithms *Reduce* and *Extend* are easy to implement. Assume now that algorithm $MetricBridgePartition$ produces an output x, y, z, t, K when applied on a metric space (M, d). This means that there is a bridge (h_{xyz}, h_{xzt}) linking K with $L = M\backslash K$ in all optimal realizations of (M, d). According to the proof of Theorem 3, such an optimal realization G can be obtained as follows.

- Compute $f(i) = d(z, i) - \frac{1}{2}(d(x, z) + d(y, z) - d(x, y))$ for all $i \in K$, and $g(i) = d(x, i) - \frac{1}{2}(d(x, z) + d(x, t) - d(z, t))$ for all $i \in L$.
- Construct a metric space $(K', d_{K'})$ as follows: if there is an element $i \in K$ with $f(i) = 0$ then set $K' = K$ and $u = i$, and define $d_{K'} = d|_K$; else build K' by adding an auxiliary element u to K, and define $d_{K'}(i, j) = d(i, j)$ for all $i, j \in K$ and $d_{K'}(i, u) = f(i)$ for all $i \in K$.

– Construct a metric space $(L', d_{L'})$ as follows: if there is an element $i \in L$ with $g(i) = 0$ then set $L' = L$ and $v = i$, and define $d_{L'} = d|_L$; else build L' by adding an auxiliary element v to L, and define $d_{L'}(i,j) = d(i,j)$ for all $i, j \in L$ and $d_{L'}(i,v) = g(i)$ for all $i \in L$.
– Construct two optimal realizations $G_{K'}$ and $G_{L'}$ of $(K', d_{K'})$ and $(L', d_{L'})$.
– Construct an optimal realization G of (M,d) by linking $G_{K'}$ and $G_{L'}$ with an edge (u,v) of length $s_{xzyt} - s_{xyzt}$.

Algorithm *OptimalRealization* uses $MetricBridgePartition$ recursively to build an optimal realization of any finite metric space (M,d). Figure 4 illustrates its use on the example of Figure 1. The possible outputs (up to symmetry) of $MetricBridgePartition$ applied on (M,d) are

Algorithm 2 *OptimalRealization*

Require: A finite metric space (M,d);
Ensure: An optimal realization G of (M,d);

if (M,d) is reducible **then**
  Apply *Reduce* on (M,d) to build an irreducible metric space (M',d');
else
  $(M',d') \leftarrow (M,d)$;
end if

Apply $MetricBridgePartition$ on (M',d');
if the output indicates that no optimal realization of (M',d') has a bridge **then**
  Apply *NoBridge* on (M',d') to build an optimal realization G' of (M',d');
else
  Let x, y, z, t, K be the output of $MetricBridgePartition$;
  Build the metric spaces $(K', d_{K'})$ and $(L', d_{L'})$ as explained above;
  Get $G_{K'}$ and $G_{L'}$ by applying *OptimalRealization* on $(K', d_{K'})$ and $(L', d_{L'})$;
  Build G' by linking $G_{K'}$ and $G_{L'}$ with an edge (u,v) of length $s_{xzyt} - s_{xyzt}$;
end if

if $(M,d) \neq (M',d')$ **then**
  Apply *Extend* to G' to build an optimal realization G of (M,d);
else
  $G \leftarrow G'$.
end if

– $x = 1, y = 3, z = 4, t \in \{6,7,8,9,10,11\}, K = \{1,2,3\}$;
– $x \in \{1,2,3\}, y = 5, z = 6, t \in \{7,8,9,10,11\}, K = \{1,2,3,4,5\}$;
– $x \in \{1,2,3,4,5\}, y = 6, z = 7, t \in \{9,10,11\}, K = \{1,2,3,4,5,6\}$.

Assume the algorithm produces the output $x = 1, y = 3, z = 4, t = 6, K = \{1,2,3\}$. Since $f(1) = 1, f(2) = 2$, and $f(3) = 1$, we construct a metric space $\mathbf{M}_1$ on $\{1,2,3,u\}$. Algorithm $MetricBridgePartition$ applied on $\mathbf{M}_1$ produces a message indicating that no optimal realization of $\mathbf{M}_1$ contains a bridge. An optimal realization G_1 of $\mathbf{M}_1$ is therefore obtained by applying the $NoBridge$ algorithm. Since $g(4) = 1, g(5) = g(6) = 2, g(7) = 4, g(8) = g(9) = g(11) = 5, g(10) = 6$, we construct a metric space $\mathbf{M}_2$ on $\{4,5,6,7,8,9,10,11,v\}$. Then, the possible outputs (up to symmetry) of $MetricBridgePartition$ applied on $\mathbf{M}_2$ are

- $x = v, y = 5, z = 6, t \in \{7, 8, 9, 10, 11\}, K = \{4, 5, v\}$;
- $x \in \{4, 5, v\}, y = 6, z = 7, t \in \{9, 10, 11\}, K = \{4, 5, 6, v\}$.

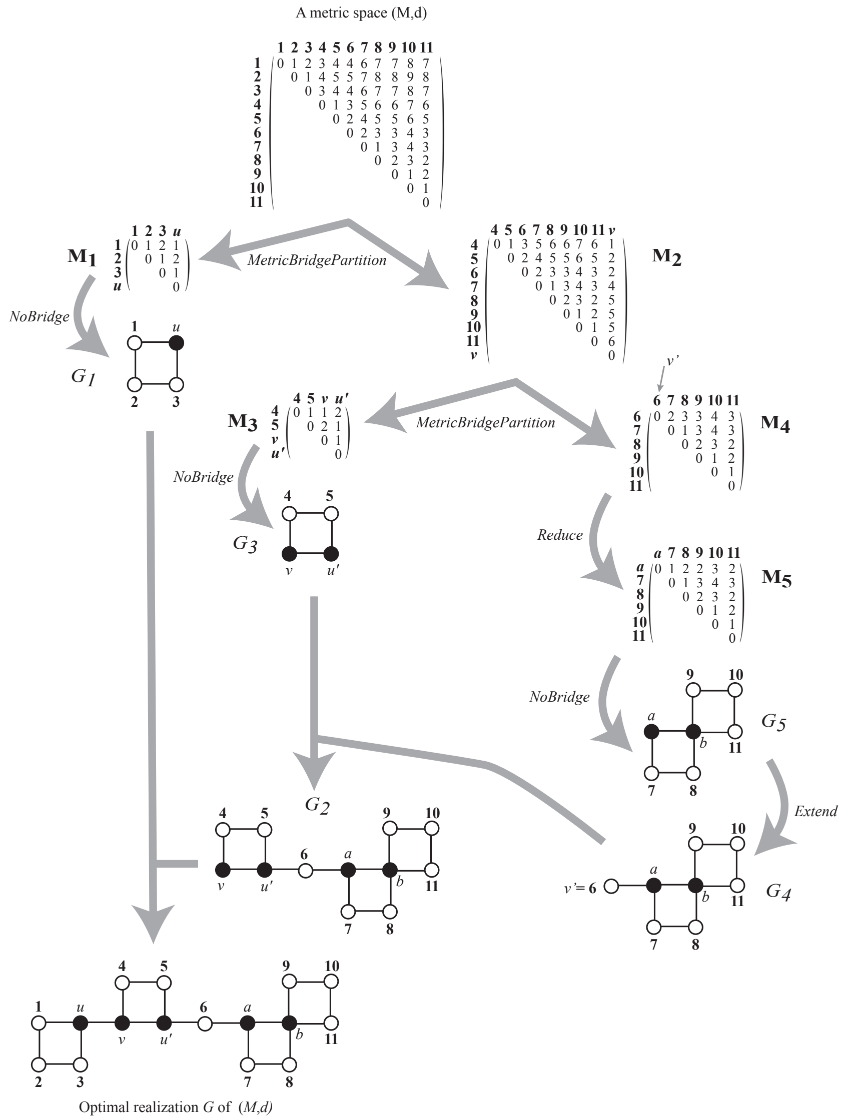


Figure 4. Construction of an optimal realization

Assume the output is $x = v, y = 5, z = 6, t = 7, K = \{4, 5, v\}$.

- Since $f(4) = 2, f(5) = 1$, and $f(v) = 1$, we construct a metric space $\mathbf{M}_3$ on $\{4, 5, v, u'\}$. Since *MetricBridgePartition* detects that no optimal realization of $\mathbf{M}_3$ has a bridge, we apply *NoBridge* on $\mathbf{M}_3$ to get an optimal realization G_3.
- Since $g(6) = 0$, we consider the metric space $\mathbf{M}_4$ induced on $\{6, 7, \cdots, 11\}$ and set $v' = 6$. $\mathbf{M}_4$ is first reduced to a metric space $\mathbf{M}_5$, where an auxiliary element a

replaces element 6. An optimal realization G_5 of $\mathbf{M}_5$ is then obtained by applying *NoBridge* (since G_5 has no bridge), and an optimal realization of $\mathbf{M}_4$ is then obtained by applying *Extend* on G_5.

Finally, G_3 and G_4 are linked together with an edge $(u', v' = 6)$ of length 1 to produce an optimal realization G_2 of $\mathbf{M}_2$; G_1 and G_2 are linked together with an edge (u, v) of length 1 to produce an optimal realization G of the original metric space (M, d).

4 Final Remarks and Conclusion

We have proved that the Metric Bridge Partition Problem is polynomially solvable. The proposed algorithm can be used to decompose any metric space (M, d) into metric spaces $(M_1, d_1), \cdots, (M_r, d_r)$ such that no optimal realization of (M_i, d_i) $(i = 1, \cdots, r)$ has a bridge. An optimal realization of (M, d) can then easily be obtained by adding some edges linking optimal realizations of $(M_1, d_1), \cdots, (M_r, d_r)$.

Although the time complexity of the decomposition algorithm is relatively high $O(|M|^6)$, this time denotes the worst possible case. In practice, some strategies can be tried in order to find as quickly as possible the right decomposition. For example, one might choose x and y by decreasing distances, as well as z and t. Notice also that in real data, there is often some noise. An open question is to find an adaptation of our decomposition algorithm in case of noisy data.

An ideal algorithm, as indicated in Theorem 1, should decompose a metric space into blocks (i.e., maximal two-connected subgraphs or bridges). The proposed algorithm is not able to detect cutpoints that do not belong to a bridge. For example, we have not been able to further decompose $\mathbf{M}_5$ in the example of Figure 4, while its optimal realization G_5 has two blocks sharing the cutpoint b. Our algorithm for the solution of the Metric Bridge Partition Problem relies on the fact that if there is a bridge (u, v) linking K and L, it is possible to decide if an element of M belongs to K or L by computing its distance to u and v. We do not know how to make such a partition using only a cutpoint u. Future work will consist in studying the more general Metric Cutpoint Partition Problem, which is to determine if the elements of a metric space (M, d) can be partitioned into two nonempty subsets K and L such that all optimal realizations of (M, d) contain a cutpoint linking K with L. The complexity of this problem is still unknown.

References

Chung, F., Garrett, M., Graham, R. and Shallcross, D.(2001) 'Distance realization problems with applications to Internet tomography'. *Journal of Computer and System Sciences* **63**:432–448.

Cunningham, J. A. (1978). 'Free trees and bidirectional trees as representations of psychological distance'. *J. Math. Psychol* **17**:165–188.

Dress, A. W. M. (1984). 'Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups : a note on combinatorial properties of metric spaces'. *Adv. in Math.* **53**:321–402.

Hakimi, S. L. and Yau, S. S. (1964). 'Distance matrix of a graph and its realizability'. *Quart. Appl. Math.* **22**:305–317.

IMRICH, W., SIMÕES-PEREIRA, J. M. S. and ZAMFIRESCU, C. M. (1984). 'On optimal embeddings of metrics in graphs'. *J. Combin. theory* **36B**:1–15.
IMRICH, W. and STOCKIǏ, E.(1972). 'The optimal embeddings of metrics into graphs'. *Sibirsk. Mat. Ž.* **13**:558–565.
LANDRY, P. A., LAPOINTE, F. J. and KIRSCH, J. A. W. (1996). 'Estimating phylogenies from distance matrices: additive is superior to ultrametric estimation'. *Molecular Biology and Evolution* **13**:818–823.
LI, M., CHEN, X., MA, B. and VITÁNYI, P.M.B. (2004). 'The similarity metric'. *IEEE Transactions on Information Theory* **50**(12):3250–3264.
MAKARENKOV, V.(2002). 'Comparison of four methods for inferring phylogenetic trees from incomplete dissimilarity matrices'. In *Classification, Clustering, and Data Analysis*, pp. 371–378. Springer, Cracow, Poland.
WINKLER, P. (1988). 'The complexity of metric realisation'. *SIAM J. Disc. math.* **1**(4):552–559.

Appendix

In this appendix we prove Theorem 3 and Theorem 4. We start with a sufficient condition for the existence of a bridge in all optimal realizations of a finite metric space (M, d). It is a corollary of Theorem 2.

Corollary 1 *Suppose (M, d) is a finite metric space to which there exist a partition of M into two non-empty subsets K, L and two different mappings $f : M \to I\!R^+$ and $g : M \to I\!R^+$ such that both (K, L, f) and (K, L, g) are nice triplets. Then every optimal realization G of (M, d) has a bridge.*

Proof Let (K, L, f) and (K, L, g) be two nice triplets with $f \neq g$, and let G be any optimal realization of (M, d). We know from Theorem 2 that all chains linking K with L go through two points c and c' such that $d^G(x, c) = f(x)$ and $d^G(x, c') = g(x)\ \ \forall x \in M$. Since $f \neq g$, we conclude that $c \neq c'$, which means that all chains linking K with L traverse a bridge containing points c and c'. $\square$

The next Theorem also provides a sufficient condition for the existence of a bridge in every optimal realization of a finite metric space (M, d).

Theorem 3 *Suppose (M, d) is a finite metric space to which there exists a partition of M into two non-empty subsets K, L with $|K| > 1$ and $|L| > 1$, and assume*
(1) $s_{ijk\ell} < s_{ikj\ell} = s_{i\ell jk}\quad \forall i, j \in K$ *and* $k, \ell \in L$.
(2) $\exists\, x, y \in K$ *and* $z, t \in L$ *such that*
$s_{xzyt} - s_{xyzt} \leq s_{ikj\ell} - s_{ijk\ell}\quad \forall i, j \in K$ *and* $k, \ell \in L$
Then every optimal realization of (M, d) has a bridge (h_{xyz}, h_{xzt}) linking K with L.

Proof Notice first that we know from (1) that the optimal realization of the metric space induced by four elements $i, j \in K$ and $k, \ell \in L$ is a tree U with $d^U(h_{ijk}, h_{ik\ell}) = s_{ikj\ell} - s_{ijk\ell} = s_{i\ell jk} - s_{ijk\ell} > 0$ (see Section 2). Let T be the optimal realization of the metric space induced by x, y, z and t, and define

$$f(i) = \begin{cases} d(z,i) - d^T(z, h_{xyz}) & \text{if } i \in K \\ d(x,i) - d^T(x, h_{xyz}) & \text{if } i \in L \end{cases}$$

and

$$g(i) = \begin{cases} d(z,i) - d^T(z, h_{xzt}) & \text{if } i \in K \\ d(x,i) - d^T(x, h_{xzt}) & \text{if } i \in L \end{cases}$$

Consider any element $i \neq x$ in K, and let U denote the optimal realization of the metric space induced on x, z, t and i. By (2), we have $d^U(h_{xiz}, h_{xzt}) \geq d^T(h_{xyz}, h_{xzt})$, and since $d^U(h_{xzt}, z) = d^T(h_{xzt}, z)$, we have

$$\begin{aligned} f(i) &= d(z,i) - d^T(z, h_{xyz}) \\ &= d^U(z, h_{xzt}) + d^U(h_{xzt}, h_{xiz}) + d^U(h_{xiz}, i) - d^T(z, h_{xyz}) \\ &\geq d^T(z, h_{xzt}) + d^T(h_{xzt}, h_{xyz}) + d^U(h_{xiz}, i) - d^T(z, h_{xyz}) \\ &= d^U(h_{xiz}, i) \geq 0. \end{aligned}$$

Since $f(x) = d(z,x) - d^T(z, h_{xyz}) = d^T(x, h_{xyz}) \geq 0$, we have $f(i) \geq 0$ for all $i \in K$. Consider now any element $i \neq z$ in L, and let U denote the optimal realization of the metric space induced on x, y, z and i. Again, $d^U(h_{xyz}, h_{xzi}) \geq d^T(h_{xyz}, h_{xzt})$ and $d^U(x, h_{xyz}) = d^T(x, h_{xyz})$. Hence,

$$\begin{aligned} f(i) &= d(x,i) - d^T(x, h_{xyz}) \\ &= d^U(x, h_{xyz}) + d^U(h_{xyz}, h_{xzi}) + d^U(h_{xzi}, i) - d^T(x, h_{xyz}) \\ &\geq d^T(x, h_{xyz}) + d^T(h_{xyz}, h_{xzt}) + d^U(h_{xzi}, i) - d^T(x, h_{xyz}) \\ &= d^T(h_{xyz}, h_{xzt}) + d^U(h_{xzi}, i) > d^U(h_{xzi}, i) \geq 0. \end{aligned}$$

Since $f(z) = d(x,z) - d^T(x, h_{xyz}) = d^T(z, h_{xyz}) > 0$, we have $f(i) > 0$ for all $i \in L$. Consider now two elements $i \in K$ and $j \in L$. We have

$$\begin{aligned} f(i) + f(j) &= d(z,i) - d^T(z, h_{xyz}) + d(x,j) - d^T(x, h_{xyz}) \\ &= d(z,i) + d(x,j) - d(x,z). \end{aligned}$$

It follows that if $i = x$ or/and $j = z$ then $f(i) + f(j) = d(i,j)$. Otherwise, let U denote the optimal realization of the metric space induced by x, z, i and j. We have $d(z,i) + d(x,j) - d(x,z) = d^U(z,i) + d^U(x,j) - d^U(x,z) = d^U(i,j) = d(i,j)$. We conclude that $f(i) + f(j) = d(i,j)$ for all $i \in K$ and $j \in L$.

We know from (1) that $h_{xyz} = h_{xyi}$ for all $i \in K$, and $h_{xzt} = h_{izt}$ for all $i \in L$. Consider now two elements i and j in L, and let U denote the optimal realization of the metric space induced by x, y, i and j. We have

$$\begin{aligned} f(i) + f(j) &= d(x,i) + d(x,j) - 2d^T(x, h_{xyz}) \\ &= d^U(x,i) + d^U(x,j) - 2d^U(x, h_{xyi}) \\ &= d^U(i,j) + 2d^U(h_{xyi}, h_{xij}) > d^U(i,j) = d(i,j). \end{aligned}$$

Consider finally two elements i and j in K, and let U denote the optimal realization of the metric space induced by i, j, z and t. Since $d^U(h_{ijz}, h_{izt}) \geq d^T(h_{xyz}, h_{xzt})$ and $d^U(h_{izt}, z) = d^T(h_{xzt}, z)$, we have

$$\begin{aligned} f(i) + f(j) &= d(z,i) + d(z,j) - 2d^T(z, h_{xyz}) \\ &= d^U(i,j) + 2d^U(h_{ijz}, h_{izt}) + 2d^U(h_{izt}, z) - 2d^T(z, h_{xyz}) \\ &\geq d(i,j) + 2d^T(h_{xyz}, h_{xzt}) + 2d^T(h_{xzt}, z) - 2d^T(z, h_{xyz}) = d(i,j). \end{aligned}$$

Since $0 < d(x,y) \leq f(x) + f(y)$ we know that $f(x)$ or/and $f(y)$ is strictly positive. We can therefore conclude that (K, L, f) is a nice triplet. The proof that (K, L, g) is a nice triplet is similar and can be obtained by permuting the roles of x, y and K with those of z, t and L.

Notice that $f \neq g$ since

$$\begin{aligned} g(i) &= f(i) + d^T(z, h_{xyz}) - d^T(z, h_{xzt}) = f(i) + d^T(h_{xyz}, h_{xzt}) > f(i) \ \ \forall i \in K \\ g(i) &= f(i) + d^T(x, h_{xyz}) - d^T(x, h_{xzt}) = f(i) - d^T(h_{xyz}, h_{xzt}) < f(i) \ \ \forall i \in L. \end{aligned}$$

By Corollary 1, we know that each realization G of (M, d) has a bridge (u, v) linking K with L. It follows from Theorem 2 that $d^G(i, u) = f(i)$ and $d^G(i, v) = g(i)$ for all $i \in M$. Since $f(i) = d^T(i, h_{xyz})$ and $g(i) = d^T(i, h_{xzt})$ for $i = x, y, z$, we conclude that $u = h_{xyz}$ and $v = h_{xzt}$. $\square$

We now give a necessary condition for the existence of a bridge.

Theorem 4 *Suppose (M,d) is an irreducible finite metric space. If there is a partition of M into two non-empty subsets K, L such that all optimal realizations of (M,d) contain a bridge (u,v) linking K with L, then*
(1) $|K| > 1$ *and* $|L| > 1$,
(2) $s_{ijk\ell} < s_{ikj\ell} = s_{i\ell jk} \quad \forall i, j \in K$ *and* $k, \ell \in L$,
(3) $\exists x, y \in K$ *and* $z, t \in L$ *such that*
•$s_{xzyt} - s_{xyzt} \leq s_{ikj\ell} - s_{ijk\ell} \quad \forall i, j \in K$ *and* $k, \ell \in L$
•$i \in K \Leftrightarrow d(x,i) - d(z,i) \leq d(x,y) - d(z,y)$.

Proof Consider a partition of M into two non-empty subsets K, L such that all optimal realizations of (M,d) contain a bridge (u,v) linking K with L. If $|K| = 1$ then the unique element in K is a vertex of degree 1 in all optimal realizations of (M,d). But this is impossible since (M,d) is irreducible. Hence $|K| > 1$, and $|L| > 1$ by symmetry.

Consider now any four elements $i, j \in K$ and $k, \ell \in L$ and let G be any optimal realization of (M,d). Since all chains linking K with L in G traverse the bridge (u,v), we have

$$\begin{aligned} s_{ikj\ell} &= d(i,k) + d(j,\ell) \\ &= d^G(i,u) + d^G(u,v) + d^G(v,k) + d^G(j,u) + d^G(u,v) + d^G(v,\ell) \\ &= d(i,\ell) + d(j,k) = s_{i\ell jk} \\ &> d^G(i,u) + d^G(j,u) + d^G(v,k) + d^G(v,\ell) \\ &\geq d^G(i,j) + d^G(k,\ell) = d(i,j) + d(k,\ell) = s_{ijk\ell}. \end{aligned}$$

Consider now four elements x, y in K and z, t in L such that $s_{xzyt} - s_{xyzt} \leq s_{ikj\ell} - s_{ijk\ell}$ for all i, j in K and k, ℓ in L, and let T be the optimal realization of the metric space induced on x, y, z and t. Also, consider any $i \in M$. If $i = x$, then $d(x,i) - d(z,i) = -d(z,x) \leq d(x,y) - d(z,y)$, and if $i = z$, then $d(x,i) - d(z,i) = d(x,z) > d(x,y) - d(z,y)$. So assume $i \neq x, z$, and let W be the optimal realization of the metric space induced on x, z, i.

- If $i \in K$, then let U be the optimal realization of the metric space induced on x, z, t and i. Since $d^U(h_{xiz}, h_{xzt}) \geq d^T(h_{xyz}, h_{xzt})$ and $d^U(h_{xzt}, z) = d^T(h_{xzt}, z)$, we have

$$\begin{aligned} d^W(x, h_{xiz}) &= d^U(x, h_{xiz}) = d(x,z) - d^U(h_{xiz}, h_{xzt}) - d^U(h_{xzt}, z) \\ &\leq d(x,z) - d^T(h_{xyz}, h_{xzt}) - d^T(h_{xzt}, z) = d^T(x, h_{xyz}). \end{aligned}$$

- If $i \in L$, then let U be the optimal realization of the metric induced by x, y, z and i. We have

$$\begin{aligned} d^W(x, h_{xiz}) &= d^U(x, h_{xiz}) = d^U(x, h_{xyz}) + d^U(h_{xyz}, h_{xzi}) \\ &= d^T(x, h_{xyz}) + d^U(h_{xyz}, h_{xzi}) > d^T(x, h_{xyz}). \end{aligned}$$

We therefore conclude that

$$\begin{aligned} i \in K &\Leftrightarrow d^W(x, h_{xiz}) \leq d^T(x, h_{xyz}) \\ &\Leftrightarrow \tfrac{1}{2}(d(x,z) + d(x,i) - d(z,i)) \leq \tfrac{1}{2}(d(x,z) + d(x,y) - d(z,y)) \\ &\Leftrightarrow d(x,i) - d(z,i) \leq d(x,y) - d(z,y). \quad \square \end{aligned}$$