Alternative Axiomatic Constructions for Hierarchical Clustering of Asymmetric Networks

Gunnar Carlsson*, Facundo Mémoli†, Alejandro Ribeiro‡ and Santiago Segarra‡
*Department of Mathematics, Stanford University
† School of Computer Science, University of Adelaide
‡Department of Electrical and Systems Engineering, University of Pennsylvania

***Abstract*—The authors have introduced an axiomatic construction for hierarchical clustering of asymmetric – i.e. weighted and directed – networks. In such construction, nodes in a two-node network cluster together at the largest of the two dissimilarities. This paper introduces two alternative constructions requiring clustering at the smallest dissimilarity and being agnostic at whether the minimum or maximum is the proper choice. Within the first framework, unilateral clustering is defined and shown to be the unique method that satisfies the proposed axioms. Within the second framework, uniform bounds are established in the minimum and maximum resolution at which clusters are formed. Unilateral clustering is used to study internal migration in the United States.**

***Index Terms*—Clustering, asymmetric networks**

I. Introduction

In asymmetric – i.e., weighted and directed – networks, the notion of proximity between nodes is not well defined since nodes can be close in one direction but far away in the other. This hinders intuitive understanding of clustering and explains why there is a large number of clustering methods for finite metric spaces – see, e.g., [1] – while methods for asymmetric networks are rarer [2]–[4]. To overcome this impediment and motivated by recent theoretical developments in clustering theory for finite metric spaces [5]–[8], we build a theory of clustering for asymmetric networks on a set of axioms that condense desirable properties of clustering methods. In this work we focus on hierarchical clustering methods whose output is a dendrogram consisting of a nested set of partitions indexed by a resolution parameter [9], [10]. Clustering methods are then maps that attribute a dendrogram to every asymmetric network. We consider the problem of designing these maps.

Our prior work in [11] introduces a first axiomatic framework for the study of hierarchical clustering of asymmetric networks. This framework is based on the axioms of value – in a two-node network the nodes merge at the lowest resolution that allows bidirectional influence – and transformation – if no pairwise dissimilarity is increased, the resolution at which clusters form cannot increase. We develop two clustering methods, reciprocal and nonreciprocal clustering, and show that every other method satisfying the aforementioned axioms lies between these two in a well-defined sense. In [12], we analyze intermediate clustering methods, i.e. contained between reciprocal and nonreciprocal clustering, that satisfy the axioms of value and transformation. Computationally tractable algorithms for all proposed clustering methods are also given

The Axiom of Value in [11], [12] insists on bidirectional influence for the formation of clusters. This requirement is well justified in some applications but in other situations unidirectional influence suffices for cluster determination. This paper develops the corresponding alternative axiomatic framework (Section III) and introduces the unilateral clustering method which is shown to be the unique hierarchical clustering method satisfying the given axioms (Section III-A). We further take an agnostic view on whether bidirectional or unidirectional influence should be enforced in the Axiom of Value, generating a second alternative axiomatic construction. We prove that in this agnostic framework any clustering method forms clusters at resolutions coarser than those of unilateral clustering and finer than those of reciprocal clustering (Section III-B). We use unilateral clustering to study the network of state-to-state migration in the United States (US) (Section IV).

Thanks to AFOSR MURI FA9550-10-1-0567, NSF CCF-1217963, and DARPA FA9550-12-1-0416.

II. Preliminaries

The network $N = (X, A_X)$ is a set of n nodes X endowed with a real valued dissimilarity function $A_X : X \times X \to \mathbb{R}_+$ defined for all pairs $x, x' \in X$. Dissimilarities $A_X(x, x')$ from x to x' are nonnegative, and null if and only if $x = x'$, but may not satisfy the triangle inequality and may be asymmetric, i.e. $A_X(x, x') \neq A_X(x', x)$ for some $x, x' \in X$. The output of hierarchically clustering the network $N = (X, A_X)$ is a dendrogram D_X, that is a nested set of partitions $D_X(\delta)$ indexed by the resolution parameter $\delta \geq 0$. Partitions in every dendrogram D_X must satisfy two boundary conditions: for the resolution parameter $\delta = 0$ each point $x \in X$ must form its own cluster, i.e., $D_X(0) = \big\{\{x\}, x \in X\big\}$, and for some sufficiently large resolution δ_0 all nodes must belong to the same cluster, i.e., $D_X(\delta_0) = \big\{X\big\}$. Partitions being nested means that if any two nodes $x, x' \in X$ are in the same block of the partition at resolution δ_0, then they stay co-clustered for all larger resolutions $\delta > \delta_0$. When x and x' are co-clustered at resolution δ in D_X we say that they are equivalent at that resolution and write $x \sim_{D_X(\delta)} x'$. For future reference, we define the two-node network $\vec{\Delta}_2(\alpha, \beta) := (\{p, q\}, A_{p,q})$ with $A_{p,q}(p, q) = \alpha$ and $A_{p,q}(q, p) = \beta$ for some $\alpha, \beta > 0$ as depicted in Fig. 1

Given a network (X, A_X) and $x, x' \in X$, a chain $C(x, x')$ is an ordered sequence of nodes in X, $C(x, x') = [x = x_0, x_1, \dots, x_{l-1}, x_l = x']$, which starts at x and finishes at x'. The *links* of a chain are the edges connecting consecutive nodes of the chain in the direction given by it. We define the *cost* of chain $C(x, x')$ as the maximum dissimilarity $\max_{i | x_i \in C(x,x')} A_X(x_i, x_{i+1})$ encountered when traversing its links in order. The directed minimum chain cost $\tilde{u}^*_X(x, x')$ between x and x' is then defined as the minimum cost among all the chains connecting x to x',

$$\tilde{u}^*_X(x, x') := \min_{C(x,x')} \max_{i | x_i \in C(x,x')} A_X(x_i, x_{i+1}). \tag{1}$$

We further define the *separation* of a network (X, A_X) as its minimum positive dissimilarity and denote it as $\text{sep}(X, A_X)$,

$$\text{sep}(X, A_X) := \min_{x \neq x'} A_X(x, x'). \tag{2}$$

An ultrametric u_X on the set X is a function $u_X : X \times X \to \mathbb{R}_+$ that satisfies symmetry $u_X(x, x') = u_X(x', x)$, identity $u_X(x, x') =$

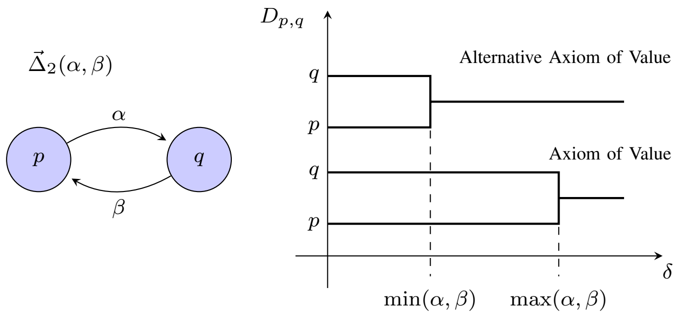

Fig. 1. Axiom of Value and Alternative Axiom of Value. For a two node network, the Axiom of Value (A1) clusters both nodes at the minimum resolution at which both can influence each other whereas the Alternative Axiom of Value (A1') clusters both nodes at the minimum resolution at which at least one of them can influence the other.

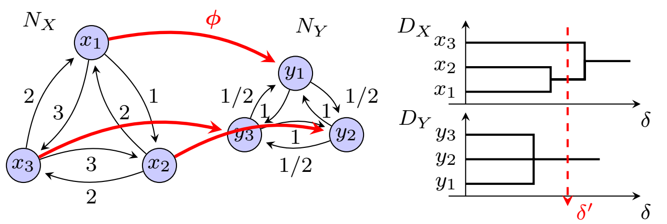

Fig. 2. Axiom of Transformation. If network N_X can be mapped to network N_Y using a dissimilarity reducing map ϕ, nodes clustered together in $D_X(\delta)$ at arbitrary resolution δ must also be clustered in $D_Y(\delta)$. For example, x_1 and x_2 are clustered together at resolution δ', therefore y_1 and y_2 must also be clustered at this resolution.

$0 \iff x = x'$ and the strong triangle inequality

$$u_X(x,x') \leq \max\left(u_X(x,x''), u_X(x'',x')\right), \tag{3}$$

for all $x, x', x'' \in X$. For a given dendrogram D_X consider the minimum resolution at which x and x' are co-clustered and define

$$u_X(x,x') := \min\left\{\delta \geq 0, x \sim_{D_X(\delta)} x'\right\}. \tag{4}$$

It can be shown that the function u_X as defined in (4) is an ultrametric on the set X, thus proving that dendrograms and finite ultrametric spaces are equivalent, [8, Theorem 9]. Ultrametrics are more convenient than dendrograms to present the results developed in this paper.

A hierarchical clustering method can be defined as a map $\mathcal{H} : \mathcal{N} \to \mathcal{D}$ from the space of networks $\mathcal{N}$ to the space of dendrograms $\mathcal{D}$, or, equivalently, as a map $\mathcal{H} : \mathcal{N} \to \mathcal{U}$ mapping every asymmetric network into the space $\mathcal{U}$ of networks with ultrametrics as dissimilarity functions. In our original axiomatic formulation, we looked for methods $\mathcal{H}$ that satisfied the following intuitive restrictions:

(A1) Axiom of Value. The ultrametric output $(X, u_{p,q}) = \mathcal{H}(\vec{\Delta}_2(\alpha,\beta))$ produced by $\mathcal{H}$ applied to the two-node network $\vec{\Delta}_2(\alpha,\beta)$, see Fig. 1, satisfies

$$u_{p,q}(p,q) = \max(\alpha,\beta). \tag{5}$$

(A2) Axiom of Transformation. Given networks $N_X = (X, A_X)$ and $N_Y = (Y, A_Y)$ and a dissimilarity reducing map $\phi : X \to Y$, that is a map ϕ such that for all $x, x' \in X$ it holds $A_X(x,x') \geq A_Y(\phi(x),\phi(x'))$; see Fig. 2. Then, the outputs $(X, u_X) = \mathcal{H}(X, A_X)$ and $(Y, u_Y) = \mathcal{H}(Y, A_Y)$ satisfy

$$u_X(x,x') \geq u_Y(\phi(x),\phi(x')). \tag{6}$$

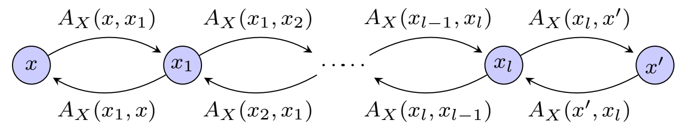

Fig. 3. Reciprocal clustering. Nodes x, x' cluster at resolution δ if they can be joined with a bidirectional chain of maximum dissimilarity δ [cf. (7)].

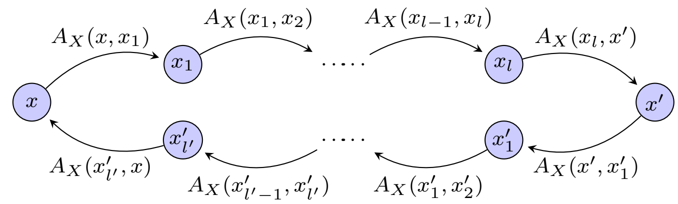

Fig. 4. Nonreciprocal clustering. Nodes x, x' cluster at resolution δ if they can be joined in both directions with possibly different chains of maximum dissimilarity δ [cf. (8)].

Axiom (A1) says that in a network with two nodes p and q, the dendrogram D_X has them merging at the maximum value of the two dissimilarities α and β. Axiom (A2) states that a contraction of the dissimilarity matrix A_X entails a contraction of the ultrametric u_X.

A hierarchical clustering method $\mathcal{H}$ is *admissible* if it satisfies axioms (A1) and (A2). Two admissible methods of interest are reciprocal and nonreciprocal clustering. The *reciprocal* clustering method $\mathcal{H}^{\text{R}}$ with output $(X, u_X^{\text{R}}) = \mathcal{H}^{\text{R}}(X, A_X)$ is the one for which the ultrametric $u_X^{\text{R}}(x,x')$ between points x and x' is given by

$$u_X^{\text{R}}(x,x') := \min_{C(x,x')} \max_{i | x_i \in C(x,x')} \bar{A}_X(x_i, x_{i+1}), \tag{7}$$

where $\bar{A}_X(x,x') := \max(A_X(x,x'), A_X(x',x))$ for all $x, x' \in X$. Intuitively, in (7) we search for chains $C(x,x')$ linking nodes x and x'. Then, for a given chain, walk from x to x' and determine the maximum dissimilarity, in either the forward or backward direction, across all links in the chain. The reciprocal ultrametric $u_X^{\text{R}}(x,x')$ is the minimum of this value across all possible chains; see Fig. 3.

Reciprocal clustering joins x to x' by going back and forth at maximum cost δ through the same chain. *Nonreciprocal* clustering $\mathcal{H}^{\text{NR}}$ permits different chains. Hence, we define the nonreciprocal ultrametric between x and x' as the maximum of the two directed minimum chain costs (1) from x to x' and x' to x

$$u_X^{\text{NR}}(x,x') := \max\Big(\tilde{u}_X^*(x,x'),\ \tilde{u}_X^*(x',x)\Big). \tag{8}$$

In (8) we implicitly consider forward chains $C(x,x')$ going from x to x' and backward chains $C(x',x)$ from x' to x. We then determine the respective maximum dissimilarities and search independently for the forward and backward chains that minimize the respective maximum dissimilarities. The nonreciprocal ultrametric $u_X^{\text{NR}}(x,x')$ is the maximum of these two minimum values; see Fig. 4.

Reciprocal and nonreciprocal clustering are of importance because they bound the range of ultrametrics generated by any other admissible method $\mathcal{H}$ in the sense stated in the following theorem.

Theorem 1 ([11]) *Consider an arbitrary network $N = (X, A_X)$ and let u_X^{R} and u_X^{NR} be the associated reciprocal and nonreciprocal ultrametrics as defined in* (7) *and* (8)*. Then, for any admissible method $\mathcal{H}$, the output ultrametric $(X, u_X) = \mathcal{H}(X, A_X)$ is such that for all pairs x, x',*

$$u_X^{\text{NR}}(x,x') \leq u_X(x,x') \leq u_X^{\text{R}}(x,x'). \tag{9}$$

In Section III we see that by modifying the admissibility criterion, i.e. by altering the axiomatic framework, the result in Theorem 1 varies.

For the Alternative Axiom of Value (A1') we obtain a uniqueness result whereas for the Agnostic Axiom of Value (A1") the admissible methods span a range of ultrametrics larger than the one in (9).

III. Alternative axiomatic constructions

Of the axioms stated in Section II, the Axiom of Value (A1) is the most open to interpretation. Requiring the two-node network in Fig. 1 to cluster at resolution $\max(\alpha, \beta)$ seems reasonable because at resolutions $\delta < \max(\alpha, \beta)$ one node can influence the other but not vice versa, which in most situations means that the nodes are different in nature and, hence, must belong to different clusters. However, it is also reasonable to accept that in some situations the two nodes should be clustered together as long as one of them is able to influence the other. To account for this possibility we replace the Axiom of Value by the following alternative.

(A1') Alternative Axiom of Value. The ultrametric output $(X, u_{p,q}) = \mathcal{H}(\vec{\Delta}_2(\alpha, \beta))$ produced by $\mathcal{H}$ applied to the two-node network $\vec{\Delta}_2(\alpha, \beta)$ satisfies

$$u_{p,q}(p, q) = \min(\alpha, \beta). \tag{10}$$

Axiom (A1') replaces the requirement of bidirectional influence in Axiom (A1) to unidirectional influence; see Fig. 1. We say that a clustering method $\mathcal{H}$ is admissible respect to the alternative axioms if it satisfies axioms (A1') and (A2).

The Alternative Property of Influence (P1) that we define next, besides its intrinsic theoretical value, is a keystone in the proof of the uniqueness result in Theorem 3.

(P1') Alternative Property of Influence. For any network $N_X = (X, A_X)$ the output ultrametric $(X, u_X) = \mathcal{H}(X, A_X)$ is such that the ultrametric value $u_X(x, x')$ between any two distinct points x and x' cannot be smaller than the separation [cf. (2)] of the network

$$u_X(x, x') \geq \text{sep}(X, A_X). \tag{11}$$

The Alternative Property of Influence (P1') states that no clusters are formed at resolutions at which there are no unidirectional influences between any pair of nodes and is consistent with the Alternative Axiom of Value (A1'). Moreover, (P1') is true if (A1') and (A2) hold as we assert in the following theorem.

Theorem 2 *If a clustering method $\mathcal{H}$ satisfies the Alternative Axiom of Value (A1') and the Axiom of Transformation (A2) then it also satisfies the Alternative Property of Influence (P1').*

Proof: See [13]. ■

In (A1') we require two-node networks to cluster at the resolution where unidirectional influence occurs. When we consider (A1') in conjunction with (A2) we can translate this requirement into a statement about clustering in arbitrary networks. Such requirement is the Alternative Property of Influence (P1') which prevents nodes to cluster at resolutions at which each node in the network is disconnected from the rest.

A. Unilateral clustering

We move on to define methods that satisfy axioms (A1')-(A2) and then bound the range of admissible methods respect to these axioms. To do so let $N = (X, A_X)$ be a given network and consider the symmetric dissimilarity function

$$\hat{A}_X(x, x') := \min(A_X(x, x'), A_X(x', x)), \tag{12}$$

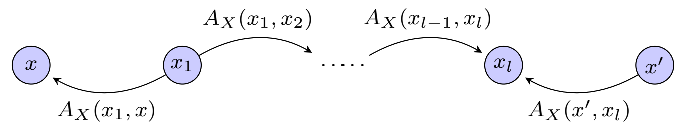

Fig. 5. Unilateral clustering. Nodes x, x' cluster at resolution δ if they can be joined with a chain with links in any direction of maximum dissimilarity δ [cf. (13)].

for all $x, x' \in X$. We define the *unilateral* clustering method $\mathcal{H}^{\text{U}}$ with output $(X, u_X^{\text{U}}) = \mathcal{H}^{\text{U}}(N)$, where u_X^{U} is defined as

$$u_X^{\text{U}}(x, x') := \min_{C(x,x')} \max_{i | x_i \in C(x,x')} \hat{A}_X(x_i, x_{i+1}), \tag{13}$$

for all $x, x' \in X$. In unilateral clustering, we first symmetrize the dissimilarities to the minimum and then look for the minimum chain cost between every pair of nodes. Equivalently, we merge two nodes at the resolution that it is possible to find a chain joining them where the directions of the links are ignored; see Fig. 5. To show that $\mathcal{H}^{\text{U}}$ is a properly defined clustering method, we need to establish that u_X^{U} as defined in (13) is a valid ultrametric. Furthermore, it can be shown that $\mathcal{H}^{\text{U}}$ satisfies axioms (A1') and (A2), as we state next.

Proposition 1 *The unilateral clustering method $\mathcal{H}^{\text{U}}$ is valid, i.e. u_X^{U} as defined in* (13) *is a valid ultrametric, and satisfies axioms (A1') and (A2).*

Proof: See [13]. ■

In the case of admissibility with respect to (A1) and (A2), we found an infinite number of clustering methods whose outcomes are uniformly bounded between those of nonreciprocal and reciprocal clustering [cf. Theorem 1]. In the case of admissibility with respect to (A1') and (A2), unilateral clustering is the unique admissible method as stated in the following theorem.

Theorem 3 *Let $\mathcal{H}$ be a hierarchical clustering method satisfying axioms (A1') and (A2). Then, $\mathcal{H} \equiv \mathcal{H}^{\text{U}}$ where $\mathcal{H}^{\text{U}}$ is the unilateral clustering method with output ultrametrics as in* (13).

Proof: See [13]. ■

Further note that in the case of symmetric networks we have $\hat{A}_X(x, x') = A_X(x, x') = A_X(x', x)$ [cf. (12)] for all $x, x' \in X$ and, as a consequence, unilateral clustering is equivalent to single linkage clustering [10, Ch. 4]. Moreover, for symmetric networks, axioms (A1') and (A2) coincide with two of the three axioms considered in [8] for finite metric spaces where single linkage was shown to be the only admissible clustering method. Thus, the result in Theorem 3 generalizes the uniqueness result in [8] since it depends on less axioms and is valid on the larger space of asymmetric networks and reduces to the uniqueness result known for single linkage clustering when considering symmetric networks.

B. Agnostic Axiom of Value

Axiom (A1) takes the position that every two-node network is clustered at $\max(\alpha, \beta)$, whereas Axiom (A1') takes the position that they should be clustered at $\min(\alpha, \beta)$. One can also be agnostic with respect to this issue and say that both of these situations are admissible. An agnostic version of axioms (A1) and (A1') is given.

(A1") Agnostic Axiom of Value. The ultrametric output $(X, u_{p,q}) = \mathcal{H}(\vec{\Delta}_2(\alpha, \beta))$ produced by $\mathcal{H}$ applied to the two-node network $\vec{\Delta}_2(\alpha, \beta)$ satisfies

$$\min(\alpha, \beta) \leq u_{p,q}(p, q) \leq \max(\alpha, \beta). \tag{14}$$

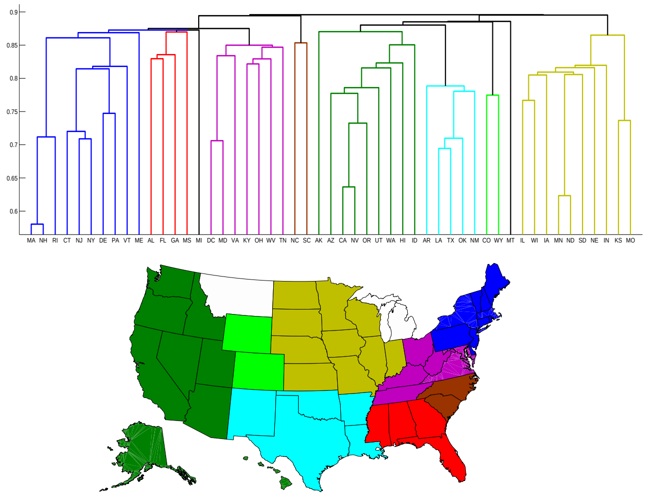

Fig. 6. State-to-state migration clustering. Dendrogram output of applying the unilateral clustering method to the network of state-to-state migration. Clusters are painted in color and identified in the map. States tend to merge with neighboring states.

Since fulfillment of (A1) or (A1') implies fulfillment of (A1"), any admissible clustering method with respect to the original axioms (A1)-(A2) or with respect to the alternative axioms (A1')-(A2) must be admissible with respect to the agnostic axioms (A1")-(A2). For methods that are admissible with respect to (A1")-(A2), we can bound the output ultrametrics as stated in the following theorem.

Theorem 4 *Consider a clustering method $\mathcal{H}$ satisfying axioms (A1") and (A2). For an arbitrary given network $N = (X, A_X)$ denote by $(X, u_X) = \mathcal{H}(X, A_X)$ the outcome of $\mathcal{H}$ applied to N. Then, for all pairs of nodes x, x'*

$$u_X^{\mathrm{U}}(x, x') \leq u_X(x, x') \leq u_X^{\mathrm{R}}(x, x'), \tag{15}$$

where u_X^{U} and u_X^{R} denote the unilateral and reciprocal ultrametrics as defined in (13) *and* (7)*, respectively.*

Proof: See [13]. ■

As in the case of the original axioms (A1)-(A2), there is an infinite number of clustering methods satisfying the agnostic axioms (A1")-(A2). By Theorem 4, given an asymmetric network (X, A_X), any hierarchical clustering method satisfying axioms (A1") and (A2) is contained between two methods. The first one, unilateral clustering, symmetrizes A_X by calculating $\hat{A}_X(x, x') = \min(A_X(x, x'), A_X(x', x))$ for all $x, x' \in X$ and then computes minimum chain costs on $(X, \hat{A}_X)$. The other method, reciprocal clustering, symmetrizes A_X by calculating $\bar{A}_X(x, x') = \max(A_X(x, x'), A_X(x', x))$ for all $x, x' \in X$ and computes minimum chain costs on $(X, \bar{A}_X)$.

IV. Numerical Experiments

We apply the unilateral clustering method $\mathcal{H}^{\mathrm{U}}$ to a network containing information about state-to-state migration in the US during year 2011 [14]. The network $N_S = (S, A_S)$ contains as a node set S every state plus Washington D.C. whereas the dissimilarity function A_S is an inverse function of the probability of state-to-state migration. More precisely, a low dissimilarity $A_S(s, s')$ for any $s, s' \in S$ indicates that an immigrant into state s' has a high probability to come from state s. Notice that N_S is inherently asymmetric. In Fig. 6 we present the dendrogram output of applying $\mathcal{H}^{\mathrm{U}}$ to N_S.

The first two states to merge in the dendrogram in Fig. 6 are Massachusetts and New Hampshire because from all the people that moved into NH, 42% came from MA, this being the highest value among the country. A similar thing occurs with individuals moving to Nevada from California and to North Dakota from Minnesota, thus the early mergings between these states in the dendrogram. If we analyze the dendrogram from top to bottom we see that the country is first divided into two clusters corresponding to an east-west division given in part by the Mississippi river. Going to lower resolutions, smaller clusters arise that we have highlighted in colors and painted the corresponding states in those colors in the US map. These clusters correspond to geographical areas where the (unidirectional) migration flow within the area is more intense than with the rest of the country. The fact that states tend to form clusters with neighboring states shows that geographical location determines the intensity of migrational flows. That is, people tend to choose nearby states as preferred destinations for internal migration.

V. Conclusion

Two alternative axiomatic frameworks for hierarchical clustering of asymmetric networks were presented. Within the first framework we defined the unilateral clustering method and showed a uniqueness result. For the second framework, we showed that unilateral and reciprocal clustering are well-defined extremes of all clustering methods satisfying the proposed axioms. Finally, we applied unilateral clustering to study state-to-state migration in the United States.

References

[1] X. Rui and D. Wunsch-II, "Survey of clustering algorithms," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 645–678, May 2005.

[2] D. Zhou, B. Scholkopf, and T. Hofmann, "Semi-supervised learning on directed graphs," *Advances in Neural Information Processing Systems*, 2005.

[3] W. Pentney and M. Meila, "Spectral clustering of biological sequence data," *Proc. Ntnl. Conf. Artificial Intel.*, 2005.

[4] M. Meila and W. Pentney, "Clustering by weighted cuts in directed graphs," *Proceedings of the 7th SIAM International Conference on Data Mining*, 2007.

[5] M. Ackerman and S. Ben-David, "Measures of clustering quality: A working set of axioms for clustering," *Proceedings of Neural Information Processing Systems*, 2008.

[6] G. Carlsson and F. Mémoli, "Classifying clustering schemes," *Foundations of Computational Mathematics*, vol. 13, no. 2, pp. 221–252, April 2013.

[7] Jon M. Kleinberg, "An impossibility theorem for clustering," in *NIPS*, Suzanna Becker, Sebastian Thrun, and Klaus Obermayer, Eds. 2002, pp. 446–453, MIT Press.

[8] Gunnar Carlsson and Facundo Mémoli, "Characterization, stability and convergence of hierarchical clustering methods," *Journal of Machine Learning Research*, vol. 11, pp. 1425–1470, 2010.

[9] G. N. Lance and W. T. Williams, "A general theory of classificatory sorting strategies 1. Hierarchical systems," *Computer Journal*, vol. 9, no. 4, pp. 373–380, Feb. 1967.

[10] Anil K. Jain and Richard C. Dubes, *Algorithms for clustering data*, Prentice Hall Advanced Reference Series. Prentice Hall Inc., Englewood Cliffs, NJ, 1988.

[11] G. Carlsson, F. Mémoli, A. Ribeiro, and S. Segarra, "Axiomatic construction of hierarchical clustering in asymmetric networks," *Proc. Int. Conf. Acoustics Speech Signal Process*, 2013.

[12] G. Carlsson, F. Mémoli, A. Ribeiro, and S. Segarra, "Hierarchical clustering methods and algorithms for asymmetric networks," *Asilomar Conf. on Signals, Systems, and Computers*, vol. (submitted), 2013.

[13] G. Carlsson, F. Mémoli, A. Ribeiro, and S. Segarra, "Axiomatic construction of hierarchical clustering in asymmetric networks," *https://fling.seas.upenn.edu/~ssegarra/wiki/index.php?n=Research.Publications*, 2012.

[14] United States Census Bureau, "State-to-state migration flows," *U.S. Department of Commerce*, 2011, Available at http://www.census.gov/hhes/migration/data/acs/state-to-state.html.