A UNIQUENESS THEOREM FOR CLUSTERING
By
Reza Bosagh Zadeh
Shai Ben-David

Akshara Dendi
FSUID – ad13r@my.fsu.edu
KLEINBERG’S AXIOMS

- Axioms
  - Scale Invariance
  - Richness
  - Consistency

- These axioms are inconsistent. That is, no algorithm can satisfy these three axioms.

- How to overcome the impossibility result? Fix $K$
Any partitioning function that satisfies *Scale Invariance, k-Richness and Consistency* is called as a *Clustering Function*.

Examples: Single-Linkage and Min-Sum

How do we distinguish between clustering functions in an axiomatic framework?

Properties are not always desirable in every clustering function.
PROPERTY: ORDER-CONSISTENCY

Order-Consistency

If two datasets \( d \) and \( d' \) have the same ordering of the distances, then for all \( k \), \( F(d, k) = F(d', k) \).
INDUCED PATH DISTANCE

- Defined as the path from $x$ to $y$, which has the smallest longest jump in it.

- Example:
PROPERTY: PATH-DISTANCE COHERENCE

➢ Path-Distance Coherence

   If two datasets \( d \) and \( d' \) have the same induced path distance then for all \( k \),
   \[
   F(d, k) = F(d', k).
   \]

➢ Certainly, this property is not expected to be hold all the time.
CHARACTERIZATION OF THE SINGLE-LINKAGE CLUSTERING

- Single-Linkage is the clustering function satisfying Order-Consistency and Path-Distance-Coherence.

- Is Path-Distance-Coherence doing all the work?

- Path-Distance-Coherence is not always desirable.
A minimum spanning tree (MST) is a spanning tree whose weight is no larger than the weight of any other spanning tree.
PROPERTY: MST-COHERENCE

MST-Coherence

If two datasets $d$ and $d'$ have the same Minimum Spanning Tree then for all $k$, $F(d, k) = F(d', k)$. 
CHARACTERIZATION OF THE SINGLE-LINKAGE CLUSTERING

- Single-Linkage is the clustering function that satisfies MST-Coherence.

- Given a weighted graph and k, Single-Linkage algorithm returns a k-partitioning by:
  • Initially, computing the MST of the induced complete graph.
  • Later, cutting the $k - 1$ most expensive edges of the MST;
  • Resulting in exactly $k$ disconnected components.

- If two distance functions give rise to the same MST then Single-Linkage will return the same clustering on both.
CHARACTERIZATION OF THE SINGLE-LINKAGE CLUSTERING

- Consistency, and k-Richness are necessary to characterize Single-Linkage

- A property X is ‘necessary’ if all remaining properties together (3 of the 4) are not enough to characterize Single-Linkage.
UNIQUENESS THEOREM

Theorem:

- Single-Linkage is the only Consistent, k-Rich, MST-Coherent, Order-Consistent partitioning function.

Proof:

- Let $F$ be any Consistent, k-Rich, MST Coherent partitioning function, and let $d$ be any distance function on $n$ points. Then, show that for all $k > 0$, $F(d, k) = SL(d, k)$.
- Consider, the partitioning $\Gamma$ such that, $SL(d, k) = \Gamma$.
- Go through a series of transformations that preserve the output.
TYPES OF EDGES

- **Outer edge:**
  Each side of the edge lies in two different clusters.

- **Inner edge:**
  Both sides of the edge lies within the cluster. There are two types of inner edges namely, redundant inner edge and non-redundant inner edge.

- **Redundant inner edge:**
  An edge that is larger than any outer edge.
Whenever we have two edges of the same type i.e. inner or outer, in neighboring positions in the edge ordering of d, we can swap their positions while maintaining the output of F by consistency.
SERIES OF TRANSFORMATIONS

- Consider some arbitrary $d$ & $k$.

- **Step 1:**
  By k-Richness of $F$, there exists a $d_1$ such that
  \[ F(d_1, k) = SL(d, k) = \Gamma. \]

- **Step 2:**
  By shrinking the inner edges of $d_1$, there exist $d_2$ such that it has same first $t$ edges as $d$. Here,
  \[ F(d_2, k) = \Gamma \text{ is maintained by Consistency.} \]
SERIES OF TRANSFORMATIONS

- **Step 3:**
  Reorder the first $t$ edges i.e. non-redundant edges and call the new dataset $d_3$. Here,
  $F(d_3, k) = \Gamma$ is maintained by consistency of partitioning function.

- **Step 4:**
  We expand all outer edges until they are larger than all inner edges and call the result $d_4$. Here,
  $F(d_4, k) = \Gamma$ is maintained by Consistency.
SERIES OF TRANSFORMATIONS

Step 5:
Reorder all outer edges until their order in relation to each other is as similar in $d$, and call the dataset $d_5$. Here, $F(d_5, k) = \Gamma$ is maintained by consistency.

Step 6:
Expand the redundant inner edges and call the new dataset $d_6$. Here, $F(d_6, k) = \Gamma$ is maintained by MST-Coherence.
SERIES OF TRANSFORMATIONS

➢ Step 7:

Change the weights of $d_6$ into exactly those of $d$ and call the new dataset as $d_7$. Here,

$F(d_7, k) = \Gamma$ is maintained by using Order-Consistency. Since we didn’t change the order of edges $d_7 = d$.

➢ 8. Thus, we have

$F(d_7, k) = F(d, k) = \Gamma = SL(d, k)$
TAXONOMY OF CLUSTERING FUNCTIONS

Overview of partitioning functions:

<table>
<thead>
<tr>
<th></th>
<th>Scale-Invariance</th>
<th>Consistency</th>
<th>k-Richness</th>
<th>MST-Coherence</th>
<th>Order-Consistency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-Linkage</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>MST cuts family</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Min-Sum k-clustering</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Constant partitioning</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Min-Sum k-clustering is Consistent, k-Rich, and Scale-invariant, but is neither Order Consistent nor MST-Coherent.
FUTURE CONSIDERATIONS

- The addition of new clustering functions and properties.

- The characterization of other clustering functions.
THANK YOU