<div align="center">CSCI 552(Spring 2021)</div>

<div align="center">Homework #2</div>

Name : Khushboo Mantri

Handout: Thursday, March 11, 2021

Due: 11:59 pm, Thursday, March 25, 2021

Total points: 35

All assignments will be submitted through Canvas. Documents will need to be in either Word or PDF format.Images need to be in jpeg format.

1. In multi-dimensional data visualization, some techniques are more suited for large volume ofdata (size) and some are better suited for data set of very high dimensions. List 2 techniquesthat are good for large size of data and 2 for very high dimensional data (there can beoverlaps). Please explain why in each case.

There are many multi-dimensional data visualization technique.

Following are the 2 technique that are good for large size of data used

1) Radial visualization

Radial graph visualization is a variation of pie chart. Radial graph shows the co-relation between parts and the whole, but it may also include sub-categories for each part of the whole. Each category in the data series plotted in a radial bar chart is assigned different color, while sub category are assigned same

2) Node – Link graph

The use of nodes and link denote connection between a group of entities is used in this type of visualization to demonstrate how things are interwired

The advantages is that it is intuitive. and good for global structure. As the complexity is $> O(N^2)$ it is not suitable for very large data set

) • The 2 technique for very high dimensional data set are below

1) Heat map.

In Heat map colored rectangle, each represent attribute. Heat map allows user to quickly grasp the state and impact of a large number of variables at once.

2) Table lens visualization
It is used to explore large amoun of tabular data. It displays regiona data in colums and rows without obscuring any data. It has ability to sort data, focous or remove unwanted data.

2. In radial layout visualization of multi-dimensional data, variables (dimensions) are places in acircle, and data points can be drawn as dots within the circle that best reflect the data points'attribute values (coordinates). This can be done through minimizing an object function todetermine the location of each dot. Derive an object function formula for ann-dimensionaldata set for this type of radial visualization.

In Radical graph visulaziation object Junction to determine the location of each dot in follows

In a circle total angle is $360°$
N is the number of attributes
∴ Angle of attribute attribute $= (2*pi)/n$

Start point $= [l_1 * \cos\theta, l_1 * \sin\theta]$

End point $= [2* \cos\theta] | 2 * \sin\theta]$

As Radial is M-dimensional, Input is in n dimensional space. Out-put is in k dimensioneel space (keen).

Object function to determine location
$$J(y) \quad \frac{\sum_{i<j} (d_{ij} - d'_{i,j})^2}{\sum_{i<j} d'_{ij}^2}$$

d_{ij} is the distance between 2 point is n-D. Space.
d_{ij} is distance between point in k-D Space.
y is data point is k-D space
location of dot

3. One benefit of radial layout of node-link visualization is that it has more space for deeperlevel nodes.

(a) Show that this benefit is still not sufficient to overcome the exponential growth of nodesin hierarchical data sets.

→31 Radial graph is not sufficient to overcome the exponential growth of nodes in hierarcical data set because hieraciral, data is not stable data set. It can grow in any where and in any direction. To plot such data on radial graph can led to not no stats to the data. It will be very hard to spot anomailes. & Radial graph is not scalable, to plot hieracal data as it is not sctable. It is a because of continous nature. As the Radial graph is not optimal the is waste of space, as the space gets wasted it is not good for exponential growthu of hierachial data set.

(b) Explain why hyperbolic tree visualization can more effectively overcome this problem.

→ 3.2. Hyperbolic tree visualization
hyperbolic tree is plot in hyperbolic space which has more spcae the eucliden space. In hyperbolic space the circumferenc of the same circle will rise exponentially & This makes tree unclutter. In this tree putting parent and children for also, will take same space if parent and children very close. As it does not take extra spau it is good for exponential data. This ability to uniformly embed. an exponential growing structure are the aspects of layout.

children's