

# CMSC 636 Data Visualization

## HomeWork 4

Anudeep Nallamothe

November 3, 2016

## 1 Introduction

This is a document about design temporal matrix visualization constructed upon EHR data set. EHR dataset is a temporal dataset which means each attribute have timestamp attached to happening of event. Following section discusses the dataset and our proposed solution to create visualization. This visualization is based on initial design by Group 5.

**Group Names:** Anudeep Nallamothe, Kishan Pawar, Harsh Vardhan Vashista, Sushanth Chaudari

**Group Number:** 2

**Group Idea:** Hierarchical Edge Bundling

## 2 Design

### 2.1 Dataset

Dataset consists of 34 columns(which we consider as dimensions). In our visualization we don't use all the columns. There are 5673 rows in total. There are 16 diseases which are described as 16 columns.

- Stress
- PTSD
- Speech
- Anxiety
- Depression
- Headache
- Sleep
- Audiology
- Vision
- Neurologic
- Alzheimer
- Cognitive
- PCS
- Endocrine
- Skull inj
- NON skull inj

All the diseases are self explanatory. So for each row if one of these diseases is 1, it represents that disease is encountered, if 0 not encountered.

## 2.2 Domain Tasks

- What are the diseases that are repeatedly encountered
- Find out the sequence of timeline of disease that each patient encountered.
- Representation of each disease so that user can identify what is what disease.

## 2.3 Mapping Task

- Our primary task is to show the occurrence of diseases in timely order(temporal). We need to show the visualization is such a way to understand the temporal dimension of data.
- We mapped the occurrence of a disease with the color and visualized its progress with a change in luminance of the color.
- Adjacency matrix is used to map patient ID to time line.

## 2.4 Encoding

- Y-axis is mapped to patientIDs and x-axis is mapped to time.
- Diseases are mapped to colors from d3 category20c library.
- The luminance of the color represents the progression of disease.

## 2.5 Critical Analysis of Current Visualization

Current visualization shows the temporal representation. We can display all the data in the same screen. Each cell represents the occurrence on a disease with color as the type of disease. This visualization doesn't have any clutter problem. Even when all the patient IDs are plotted there is no clutter problem. Complete summary of data is presented.

## 2.6 Other Proposed Visualizations

We will discuss the design and their usage of other visualizations proposed in the class.

### Group 1:

- This visualization addresses the change task.
- It shows all the encounters of a disease in one screen but for only one patient ID at a time
- This visualization doesn't use overview plus detail technique. It shows data for one patient.
- It shows the temporal changes of data. Each vertical line represents one disease and its progression.
- This visualization doesn't introduce clutter problem. Each visualization progression is shown clearly. However if all the patients are shown simultaneously then it introduces clutter.
- Yes, this visualization is visually pleasing.
- **Pros:** This visualization have temporal representation. Clearly represents each disease clearly. No cluttering problem.
- **Cons:** This visualization doesn't summarize the progression on all the patients. So because of this it is hard to observe a pattern in progression. No overview of all the patients.

### Group 2:

- This visualization addresses the change task by moving node from one cell to the other
- It shows all the encounters using different grid cells but only for one patient ID at a time.

- Doesn't use overview detail technique. It shows data for one patient at a time.
- It visualizes temporal changes for one patient ID at a time.
- This visualization doesn't introduce clutter problem for one patient ID, but when used for all the patients at one time, cluttering happens.
- It is visually pleasing.
- **Pros:** This visualization have temporal representation. Clearly represents each disease. This visualization show the progression of each disease as well. No cluttering problem for one patient.
- **Cons:** No overview of all the patients.

### Group 3:

- This visualization addresses the change task by representing each task with one plot.
- It shows all the encounters using different plots but again only for one patient ID at a time like other two visualizations.
- Doesn't use overview detail technique. It shows data for one patient at a time.
- It visualizes temporal changes for one patient ID at a time.
- This visualization doesn't introduce clutter problem for one patient ID, but when used for all the patients at one time, cluttering happens.
- It is visually pleasing.
- **Pros:** This visualization have temporal representation. Clearly represents each disease clearly. This visualization show the progression of each disease as well. No cluttering problem for one patient.
- **Cons:** No absolute values and occlusion due to overlapping of plots.

### Group 4:

- This visualization addresses the change task by representing each task with one plot.
- It shows all the encounters using different plots but again only for one patient ID at a time like other visualizations.
- Doesn't use overview detail technique. It shows data for one patient at a time.
- It visualizes temporal changes for one patient ID at a time.
- This visualization doesn't introduce clutter problem for one patient ID, but when used for all the patients at one time, cluttering happens.
- It is visually very pleasing.
- **Pros:** This visualization have temporal representation. Clearly represents each disease. This visualization show the progression of each disease as well. No cluttering problem for one patient.
- **Cons:** No absolute values and all the patient IDs are not visualized.

### Group 5:

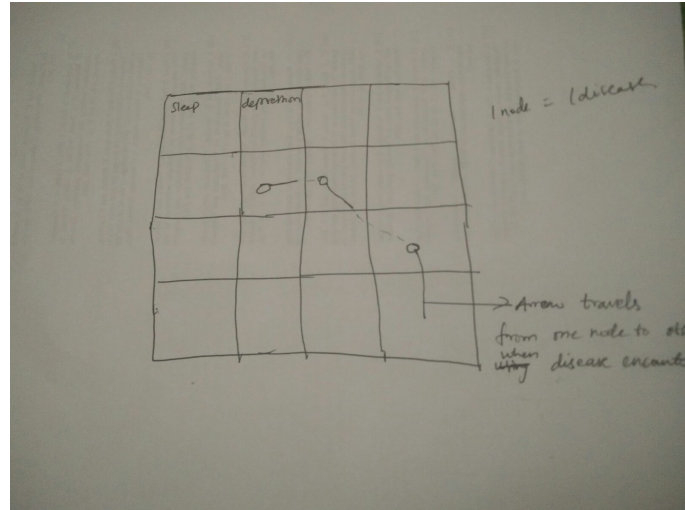


Figure 1: Sketch of Visualization

- This visualization doesn't address the change task. It only shows the number of occurrences.
- It shows all the aggregate number of encounters into one cell of the adjacency matrix. It shows the complete visualization in the screen.
- This uses overview detail technique. It shows aggregate data.
- It visualizes temporal changes for one patient ID at a time.
- This visualization doesn't introduce clutter problem.
- It is visually pleasing.
- **Pros:** This visualization don't have temporal representation. Clearly represents each disease clearly. This visualization show the aggregate of each disease.
- **Cons:** No absolute values about the timeline of each disease.