



Exploration of Street Network Node and Edge Features

CAP5771 Final Project
Quentin Goss

`quentingoss0323@floridapoly.edu`

Research Question

1. What do the features of street network edges and vertices tell us about a street network?
2. Which features are dominant?
3. How can these features be used to make a general classification of a street network?

Exploratory Data Analysis Nodes

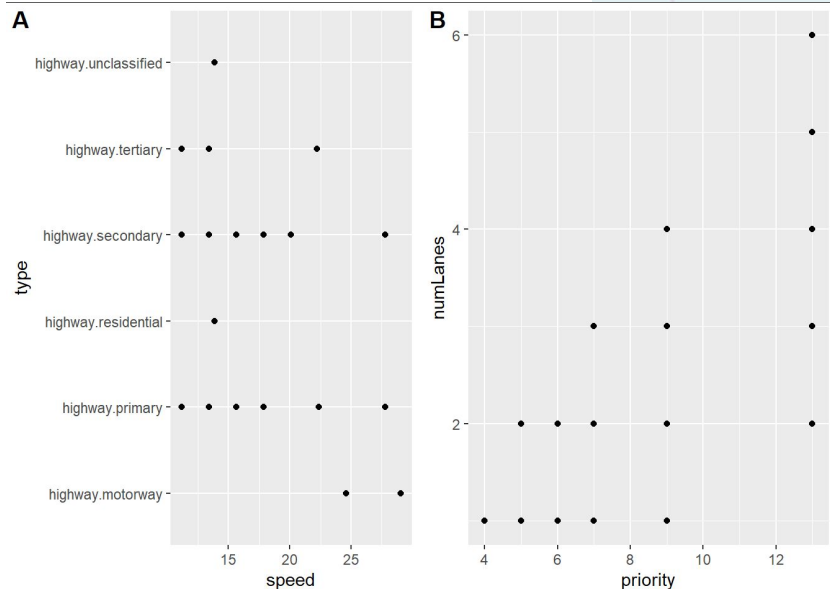
```
##          id          x          y          z
## 1201200159: 1  Min.   : 388.1  Min.   : 297.5  Mode:logical
## 1201200164: 1  1st Qu.:2020.0  1st Qu.:2134.4  NA's:3425
## 1201200194: 1  Median :2986.0  Median :3168.5
## 1201200198: 1  Mean   :2909.8  Mean   :3145.8
## 1201200223: 1  3rd Qu.:3824.2  3rd Qu.:4157.2
## 1306186991: 1  Max.   :5210.8  Max.   :5777.5
## (Other)    :3419
##          type          tlType          tl          radius
## dead_end      :    5  Mode:logical          :3265  Mode:logical
## priority      :3251  NA's:3425  joinedS 0 :    2  NA's:3425
## right_before_left:    9  1700923661:    1
## traffic_light : 160  3675134324:    1
##                               4860700048:    1
##                               542892026 :    1
##                               (Other)   : 154
## keepClear      rightOfWay
## Mode:logical    Mode:logical
## NA's:3425        NA's:3425
```

Conclusion 1: We may not apply principal component analysis to the node data, nor may we apply association rules learning to the node data because there is only one relevant categorical feature. Therefore, we do not consider node data to be considered in the classification of the SN.

Exploratory Data Analysis

Edges

```
##                                     type      numLanes      speed
## highway.residential :3893      Min.   :1.000      Min.   :11.18
## highway.tertiary     :1354      1st Qu.:1.000      1st Qu.:13.89
## highway.secondary    : 896      Median :1.000      Median :13.89
## highway.primary      : 668      Mean   :1.275      Mean   :16.17
## highway.unclassified: 214      3rd Qu.:1.000      3rd Qu.:17.88
## highway.motorway     : 49      Max.   :6.000      Max.   :29.06
## (Other)              : 0
##
##      priority      spreadType      name
## Min.   : 4.000      right :6202      Central Boulevard : 164
## 1st Qu.: 4.000      center: 872      Washington Street : 164
## Median : 4.000                                     Summerlin Avenue  : 162
## Mean   : 5.328                                     Mills Avenue      : 153
## 3rd Qu.: 6.000      Westmoreland Drive: 142
## Max.   :13.000      Amelia Street    : 139
##                                     (Other)          :6150
##
##      length
## Min.   : 0.236
## 1st Qu.: 21.064
## Median : 53.082
## Mean   : 69.526
## 3rd Qu.: 99.742
## Max.   :605.580
##
```



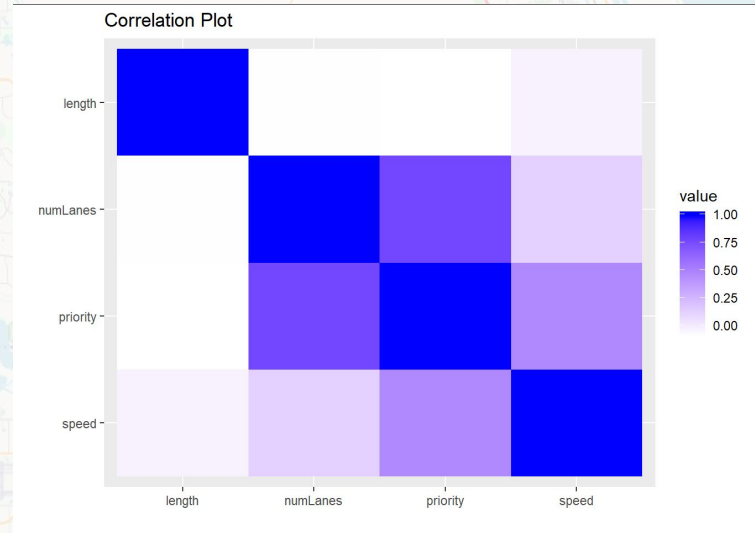
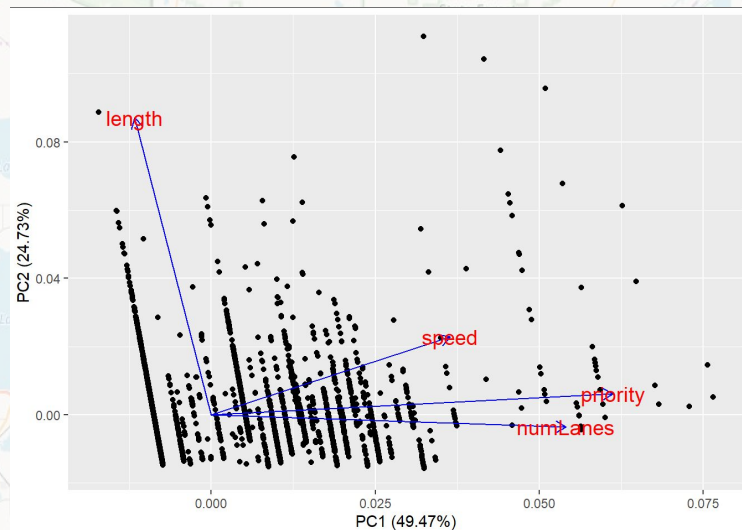
PCA: numLanes, speed, priority, length

ARL: type, numLanes, speed, priority, spreadType, names

Methods

Principal Component Analysis

```
## Importance of components:  
##               PC1      PC2      PC3      PC4  
## Standard deviation    1.4067 0.9945 0.9405 0.38431  
## Proportion of Variance 0.4947 0.2473 0.2211 0.03692  
## Cumulative Proportion 0.4947 0.7420 0.9631 1.00000
```



Conclusion 2: The dominant numerical features of an edge are **numLanes**, **priority**, and **speed**.

Methods

Association Rules Learning w/ Apriori

```
## Parameter specification:
## confidence minval smax arem aval originalSupport maxtimessupport minlen
##      0.5      0.1      1 none FALSE      TRUE      5      0.1      5

## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...372 item(s), 7074 transaction(s) done [0.00s].
## sorting and recoding items ... [12 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 5 done [0.00s].
## writing ... [10 rule(s)] done [0.00s].
## creating S4 object ... done [0.00s].

##      lhs                                     rhs      support confidence      lift count
## [2] {spreadType=right,
##      speed=22.22,
##      numLanes=1,
##      priority=6}      => {type=highway.tertiary}      0.1480068      1.0000000      5.224520      1047
## [5] {spreadType=right,
##      speed=13.89,
##      numLanes=1,
##      priority=4}      => {type=highway.residential}      0.5264348      1.0000000      1.817108      3724
```

Conclusion 3: The features **type**, **spreadType**, **speed**, **numLanes**, and **priority** may be used to create association rules through Apriori, with support = 0.1, confidence = 0.5, and minLen = 5 to produce a small number of rules that forms a general classification of the street network.

Future Work

1. Explore a variety of street networks
2. `Type=highway.residential` are often short in length and consist of many more edges than other types.
 - a. Skews the data unnaturally towards **highway.residential**
 - b. Explore ways to reduce this unnatural skew.

