What's in the network?
A stepwise overview of working with networked data in R

Tine Van Calster, Michael Reusens, María Óskarsdóttir, Sandra Mitrović, Jasmien Lismont, Jochen De Weerdt, Wilfried Lemahieu, Bart Baesens, and Jan Vanthienen

KU Leuven, Dept. of Decision Sciences and Information Management
Naamsestraat 69, B-3000, Leuven, Belgium; {GivenName.LastName@kuleuven.be}

Networks are everywhere!
• They represent any type of connection between persons or objects
• Many applications in marketing, fraud, transportation, retail, biology, research citation, etc.
• Abundance of data leads to new challenges

### Networks Everywhere!

#### Business cases

- Credit card fraud
- Social media network
- Call network

#### Analysis & Modeling

- Data structure
- Adjacency matrix
- Edge list
- Sparse matrix

#### Network visualization

- Adacency matrix
- Edge list
- Sparse matrix

#### Network learning

- Featurization
- Network learning
- Graph sampling

### Data structure

#### Adjacency matrix

<table>
<thead>
<tr>
<th></th>
<th>María</th>
<th>Michael</th>
<th>Jochen</th>
</tr>
</thead>
<tbody>
<tr>
<td>María</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Michael</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Jochen</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

#### Edge list

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>María</td>
<td>Michael</td>
<td>5</td>
</tr>
<tr>
<td>Michael</td>
<td>María</td>
<td>3</td>
</tr>
<tr>
<td>Time</td>
<td>Bart</td>
<td>6</td>
</tr>
<tr>
<td>Bart</td>
<td>Jochen</td>
<td>10</td>
</tr>
</tbody>
</table>

#### Sparse matrix

<table>
<thead>
<tr>
<th></th>
<th>María</th>
<th>Michael</th>
<th>Time</th>
<th>Bart</th>
<th>Jochen</th>
</tr>
</thead>
<tbody>
<tr>
<td>María</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Michael</td>
<td>0</td>
<td>-</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Time</td>
<td>6</td>
<td>3</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bart</td>
<td>0</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Jochen</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

#### Code Examples

```r
library(igraph)

# create graph from adjacency matrix
V(g)$name <- c("María", "Michael", "Tine", "Bart", "Jochen")
E(g)$weight <- c(5, 6, 3, etc.)
V(g)$churn <- c(1, 0, 0, 0, 1, etc.)

# calculate degree of each node
deg <- degree(g, mode="all")

# centralization measures
V(g)$page_rank <- page_rank(g, algo = "prpack", vids = V(g), directed = F, damping = 0.85, weights = NA, personalized = c(1, 0, 0, 0, 1, etc.))$vector
V(g)$betweenness <- betweenness(g, directed = F)

# create edgelist from graph
edges <- as.data.frame(get.edgelist(g))
# select columns Source, Target, Weight
colnames(edges) <- c("Source", "Target", "Weight")

# plot network
forceNetwork(Links = d3g$links, Nodes = d3g$nodes, NodeColor = "#808080", NodeID = "name", Source = "Source", Target = "Target", NodeSize = "age", colourScale = JS("d3.scaleOrdinal(d3.schemeCategory10)");,
    Group = "churn", zoom = TRUE, legend = TRUE,
    theme_graph(foreground = "steelblue", fg_text_colour = "white") + geom_node_text(aes(label = name), size = 2)
    )
```

---

**For more layouts:**

- Force-directed layout: `forceNetwork(Links = d3g$links, Nodes = d3g$nodes, NodeColor = "#808080", NodeID = "name", Source = "Source", Target = "Target", NodeSize = "age", colourScale = JS("d3.scaleOrdinal(d3.schemeCategory10)");, Group = "churn", zoom = TRUE, legend = TRUE, theme_graph(foreground = "steelblue", fg_text_colour = "white") + geom_node_text(aes(label = name), size = 2))`

- NetworkD3: `library(networkd3)`

---

**For more on sampling methods:**

- `igraph::sample_*`

---

**For more on graph visualization:**

- donut graph visualizations in web pages, markdown documents, and shiny apps

---

**Sparse matrices are useful when you have a lot of missing values take advantage of white space**

---

**Embed graph**

- `igraph::layout`
- `igraph::igraphopt`
- `igraph::igraphopt`