

Lec14 Intro to datasets

- *Zachary Karate Club* : 34 nodes and 78 edges

Lect15 Ingredients Network

(butter, sugar, flour, vanilla)->(oil, water)->(Tomato, Onion, Pepper, garlic)

Lec16 Synonymy Network

Love->Friend->..... ; Enemy->Hate->..... Love X Hate. But there is a path from Love to Hate _(REASON : Degradation of Synonymity)

Lec17 Web Graph

- Webpage linking -> GOOGLE unleashed it! RAnking of search

Lec18 Social Network Datasets

- Who could be his next friend? - Facebook

Lec19 Real Work Network Datasets

Eg. of networks

1. Friendship (undirected)
2. Road(undirected)
3. Citation(directed)
4. Email(directed)
5. Collaboration/ coauthorship (undirected)

Data Formats

- CSV
- GML
- Pajek Net
- GraphML
- GEXF (By Gephi)

csv -> .csv/.txt _

-> edge list, adj list, weighted edge list

gml_

-> Graph Modelling Language

Without attributes

```
graph [  
    node [  
        id a  
        label 'Node A'  
    ]  
    node [  
        id b  
        label 'Node b'  
    ]  
    edge [  
        source a  
        target b  
        label 'E1'  
    ]  
]
```

With attributes

```
graph [  
    heirarchic 1  
    directed 1  
    node [  
        id a  
        label 'Node A'  
        graphics[  
            x 0  
            y 0  
        ]  
    ]  
    node [  
        id b  
        label 'Node b'  
        graphics[  
            x 20  
            y 30  
        ]  
    ]  
    edge [  
        source a  
        target b  
        label 'E1'  
        LabelGraphics[  
            text 'Happy New Year'  
        ]  
    ]  
]
```

Pajek Net

- uses .net/.paj extension


mod02lec19.mp4

Pajek Net Format : Uses .NET extension

```
*Vertices 82670
1 "entity"
2 "thing"
3 "anything"
4 "something"
5 "nothing"
6 "whole"

*arcs
4244 107
4244 238
4244 4292
4247 107
4248 1
4248 54
```

Real-World Network Data Sets 11 / 16




Videos mod02lec19.mp4 Thu Aug 3 7:33 PM

Pajek Net Format

```
*Vertices 9
*Edges
1 2
1 9
2 9
2 3
2 8
3 8
3 4
4 5
4 7
5 7
5 6
6 4
```

Real-World Network Data Sets 12 / 16



mod02lec19.mp4

Pajek Net Format with attributes

```
*arcs
4244 107 5
```

14:04 4:57

Real-World Network Data Sets 13 / 16 NPTEL

GraphML

mod02lec19.mp4

GraphML Format: Uses .graphml extension

```
<?xml version="1.0" encoding="UTF-8"?>
<graphml xmlns="http://graphml.graphdrawing.org/xmlns"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns
http://graphml.graphdrawing.org/xmlns/1.0/graphml.xsd">
<graph id="G" edgedefault="undirected">
<node id="n0"/>
<node id="n1"/>
<edge id="e1" source="n0" target="n1"/>
</graph>
</graphml>
```

14:16 4:45

Real-World Network Data Sets 14 / 16 NPTEL

mod02lec19.mp4

GraphML Format with attributes

```
<?xml version="1.0" encoding="UTF-8"?>
<graphml xmlns="http://graphml.graphdrawing.org/xmlns" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns http://graphml.graphdrawing.org/xmlns" >
  <key id="d0" for="node" attr.name="color" attr.type="string">
    <default>yellow</default>
  </key>
  <key id="d1" for="edge" attr.name="weight" attr.type="double"/>
  <graph id="G" edgedefault="undirected">
    <node id="n0">
      <data key="d0">green</data>
    </node>
    <node id="n1"/>
    <node id="n2">
      <data key="d0">blue</data>
    </node>
    <node id="n3">
      <data key="d0">red</data>
    </node>
    <edge id="e0" source="n0" target="n2">
      <data key="d1">1.0</data>
    </edge>
    <edge id="e1" source="n0" target="n1">
      <data key="d1">1.0</data>
    </edge>
    <edge id="e2" source="n1" target="n3">
      <data key="d1">2.0</data>
    </edge>
  </graph>
</graphml>
```

Real-World Network Data Sets 15 / 16 NPTEL

GEXF

mod02lec19.mp4

GEXF Format : Graph Exchange XML Format

```
<?xml version="1.0" encoding="UTF-8"?>
<gexf xmlns="http://www.gexf.net/1.2draft" version="1.2">
  <meta lastmodifieddate="2009-03-20">
    <creator>Gexf.net</creator>
    <description>A hello world! file</description>
  </meta>
  <graph mode="static" defaultedgetype="directed">
    <nodes>
      <node id="0" label="Hello" />
      <node id="1" label="Word" />
    </nodes>
    <edges>
      <edge id="0" source="0" target="1" />
    </edges>
  </graph>
</gexf>
```

Real-World Network Data Sets 16 / 16 NPTEL

Lec20 How to download datasets

- Google Network datasets

(snap.stanford.edu/data, uci network data, konect, sonet for graphML datasets)