Using CINET

NSF Software Development for CyberInfrastructure Grant OCI-1032677
Additional support by grants from DTRA V&V, DTRA CNIMS, NSF NetSE, NSF DIBBS

Team
CINET: Applications

• Granite
  – Network structural analyses.

• GDS Calculator (GDSC)
  – Complete network dynamics on networks.

• EDISON
  – Forward trajectory (dynamics) on networks.
Granite: Initial Screen

• Go to:
  – http://cinet.vbi.vt.edu/granite/granite.html
  – or http://cinet.vbi.vt.edu and click Granite

• Then login

• To create a new account, click register
Features

Available features:

- Network Analysis
- Network Generators
- Network List
- Measure List
- Visualization
- NetScript
Networks and Properties

Network
- a set of nodes, representing entities, depicted by circles
- a set of edges, representing relationships, depicted by lines

A network with 6 nodes and 7 edges
Density

Number of edges / max. no. of possible edges

\[ \text{Density} = \frac{m}{\binom{n}{2}} = \frac{2m}{n(n-1)} \]

Density = \( \frac{2 \times 7}{5 \times 6} = \frac{7}{15} = 0.47 \)
A **triangle** consists of three nodes with an edge between each pair of nodes. (It is also called a 3-clique.)

\[ \{u, v, w\} \] is a triangle.

2 triangles in this network.
Connectivity:

A network is connected if there is a path between every pair of nodes.

A path between u and v is a sequence of edges that you can follow to go from node u to node v.
Diameter

A path between $u$ and $v$ is a sequence of edges that you can follow to go from node $u$ to node $v$.

Path length: number of edges in the path

The red path from $u$ to $v$ has length 3.

Diameter

= max (length of shortest path between any pair of nodes).

The diameter of the example network is 3.
Network Analysis

- In the menu bar, select network analysis
- You can see a list of analyses done earlier
- To perform a new analysis, click +New Analysis
- Type a name for the analysis
- Select one or more networks
  - You can browse or use the search box
- You can see the list of selected networks
- Click Continue
Network Analysis (cont.)

- **Select** one or more **measures**
  - You can browse or use the search box
  - You can see some details of the measures
  - If necessary, provide parameter values
- You can see the list of selected measures
- Click **Analyze**
- The new analysis is now in the list
- Look at the **status**
- When it is **COMPLETED**, click **View Report**.
- See the results in the report section
- To download the results, click **Download**.
Exercise I
Brief Discussion of Exercise I(a)

**Purpose:** To quickly familiarize attendees with the use of CINET through the Granite interface.

**Short Description:** For each of the two networks, namely *Dolphins Social Network in NZ* and the *Erdős collaboration network*, use CINET to find the following:

- density,
- number of triangles, and
- diameter.

**Procedure:** Discussed in the handout.
Solution to Exercise I(a)

Notes:

- In the table below, $G_1$ denotes the **Dolphins Social Network** and $G_2$ denotes the **Erdős Collaboration Network**.
- For simplicity, the density values have been rounded to six decimal places.

<table>
<thead>
<tr>
<th></th>
<th>#Nodes</th>
<th>#Edges</th>
<th>Density</th>
<th>#Triangles</th>
<th>Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_1$</td>
<td>62</td>
<td>159</td>
<td>0.084082</td>
<td>95</td>
<td>8</td>
</tr>
<tr>
<td>$G_2$</td>
<td>6927</td>
<td>11,850</td>
<td>0.000494</td>
<td>5973</td>
<td>4</td>
</tr>
</tbody>
</table>
Purpose: Use CINET to check whether there is any correlation between pairs of graph parameters.

Note: Background in programming will be helpful in doing this exercise.

Short Description:

- Please use the five networks specified in the handout for this exercise. (Let $G_1$ through $G_5$ denote these networks.)
- The handout shows the average node degree for each of the networks. (Let $\Delta_1$ through $\Delta_5$ denote these values.)
- Find the number of triangles in each of the networks. (Let $T_1$ through $T_5$ denote these values.)
Brief Discussion of Exercise I(b)

- Find the **diameter** of each network. (Let $D_1$ through $D_5$ denote these values.)
- Find the **Pearson Correlation Coefficient** (PCC) $r_1$ for the sample $\{(\Delta_1, T_1), \ldots, (\Delta_5, T_5)\}$. (Please see the handout for the definition and significance of PCC.)
- Find the PCC $r_2$ for the sample $\{(\Delta_1, D_1), \ldots, (\Delta_5, D_5)\}$.

**Notes:**

- It will also be useful to prepare scatter plots of the values computed. (See handout for details.)
- You may also want to study correlations between other pairs of graph parameters.
Solution to Exercise I(b)

<table>
<thead>
<tr>
<th></th>
<th>#Nodes</th>
<th>#Edges</th>
<th>Avg. degree ($\Delta$)</th>
<th>#Triangles ($T$)</th>
<th>Diameter ($D$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_1$</td>
<td>10,729</td>
<td>11,000</td>
<td>2.051</td>
<td>15,834</td>
<td>12</td>
</tr>
<tr>
<td>$G_2$</td>
<td>6927</td>
<td>11,850</td>
<td>3.421</td>
<td>5973</td>
<td>4</td>
</tr>
<tr>
<td>$G_3$</td>
<td>10,670</td>
<td>22,002</td>
<td>4.124</td>
<td>17,144</td>
<td>10</td>
</tr>
<tr>
<td>$G_4$</td>
<td>10,900</td>
<td>31,180</td>
<td>5.721</td>
<td>82,856</td>
<td>9</td>
</tr>
<tr>
<td>$G_5$</td>
<td>33,696</td>
<td>180,811</td>
<td>10.732</td>
<td>725,311</td>
<td>13</td>
</tr>
</tbody>
</table>

Pearson Correlation Coefficient (PCC) Values:

- PCC value $r_1$ ($\Delta$ vs $T$) = 0.946081
- PCC value $r_2$ ($\Delta$ vs $D$) = 0.434508
Exercise 1(b): Avg. Degree vs #Triangle

![Graph showing the relationship between average degree and number of triangles]
Exercise I(b): Avg. Degree vs Diameter
Network Generation
Random Networks

- Edges are added randomly

Erdős-Rényi, \( G(n, p) \), network

- Each potential edge is added with probability \( p \)

A \( G(n, p) \) network with \( p = 1/3 \)

A star graph: is a deterministic graph
Network Generators

- In the menu bar, select **Network Generators**
- You can see a list of generators created earlier
- Click **+New Network Generator**
- Type a **name for the generator**
- **Select** one or more **generators**
  - You can browse or use the search box
- You can see the list of selected generators
- Specify **parameters** if required and click **submit**
- Click **Generate**
Network Generators (cont.)

- The new generator is now in the list of generators
- Look at the status
- When it is COMPLETED, click View Report.
- See the results in the report section
- To download the network, click Download.
Add a New Network

- In the menu bar, select **Networks**
- Click **+New Network**
- Select **Directly upload a file**
- Click **Done**
- Click **Choose File**
- Provide a name of the network and other info
- Click **Save**
- Now you can see the added network in the list
Network Visualization

- In the menu bar, select **Networks**
- You can see the list of networks
- Click on a network name to visualize
- Click **visualization** (on the right hand side)
- Click **+Add Visualization**
- Select visualization parameters
  - leave them as they are to use the default values
- Click **Generate**
Exercise II
Connectivity:

A network is connected if there is a path between every pair of nodes.

A path between u and v is a sequence of edges that you can follow to go from node u to node v.

Connected network

Disconnected network
Each maximal group of nodes that are connected (by a path) with each other is a **component**.

- **Largest component (size: 5)**
- **3 components in this graph**

- Each largest component is also called "Giant component".
A **bridge** is an edge whose removal increases the number of components in the graph. 

(u, v) is a bridge edge.
**Purpose:** To familiarize attendees with the random graph generation facility of CINET.

**Short Description:** Generate three Erdős-Renyi random graphs, also known as $G(n, p)$ graphs, with the following parameters:

- number of nodes $n = 1000$
- edge probability $p = 0.01, 0.02, 0.05$

For the each graph, compute:

- the size of the largest component,
- the number of triangles, and
- the number of bridge edges.

**What do the results suggest?**
Solution to Exercise II(a)

**Note:** In the following table “GC Size” is used for ”Giant Component Size”.

<table>
<thead>
<tr>
<th>Prob.</th>
<th>#Nodes</th>
<th>#Edges</th>
<th>GC size</th>
<th>#Triangles</th>
<th>#Bridges</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>1000</td>
<td>5115</td>
<td>1000</td>
<td>168</td>
<td>0</td>
</tr>
<tr>
<td>0.02</td>
<td>1000</td>
<td>9952</td>
<td>1000</td>
<td>1311</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>1000</td>
<td>24,902</td>
<td>1000</td>
<td>10,295</td>
<td>0</td>
</tr>
</tbody>
</table>
**Purpose:** Experimentation with random graph generation facility of CINET.

**Short Description:** Consider $G(n, p)$ random graphs with $n = 1000$ nodes. As we increase the edge probability $p$, the number of edges in the generated graph increases. Find the smallest probability value at which the generated graph has a certain property such as:

- the largest component has as least 900 nodes,
- the graph is connected,
- the graph has no bridge edge.

**Hint:** For each property, try a binary search on the probability $p$. 
Ex. II(b): Prob. vs Giant Component Size

![Graph showing the relationship between edge probability and size of the giant component. The x-axis represents edge probability ranging from 0.005 to 0.008, and the y-axis represents the size of the giant component ranging from 992 to 1000. The graph shows a linear increase in the size of the giant component as the edge probability increases.]
Ex. II(b): Prob. vs #Bridges