**Vivekanand Education Society’s Institute of Technology**

**Department of AI &DS**

****

**Subject: Social Media Analytics**

**Class: D16ad**

| **Roll No: 30** | **Name:** [**SUHANEE KANDALKAR**](mailto:2021.suhanee.kandalkar@ves.ac.in) |
| --- | --- |
| **Practical No: 6** | **Title:** Structure Based Social Media Analytics |
| **DOP:** | **DOS:** |
| **Grades:** | **LOs Mapped:** |
| **Signature:** |  |

**Title:** Structure Based Social Media Analytics

**Aim:** To Develop structure based social media analytics model for Social Media Network to find

1. To calculate Degree centrality, Closeness centrality, Betweenness Centrality
2. To identify the most influential node.

**Theory:**

### **Network Analytics Layer of Social Media Analytics**

Social Media Analytics (SMA) consists of multiple layers that provide insights into different aspects of social media networks. The **Network Analytics Layer** focuses on the structural relationships between nodes (users, posts, hashtags) in a social media network. It helps in understanding the influence of users, detecting communities, and finding key players in a network.

The **Network Analytics Layer** includes:

1. **Nodes and Edges Representation:** Social media platforms can be represented as graphs, where nodes represent users and edges represent interactions (e.g., likes, comments, shares, or follows).
2. **Network Centrality Analysis:** Identifies key influencers in the network using centrality metrics.
3. **Community Detection:** Finds groups of closely connected nodes (clusters).
4. **Information Flow Analysis:** Examines how information propagates through the network.

The primary focus of this study is on **centrality analysis**, which helps determine the most important nodes in a network.

### **Centrality Measures**

Centrality measures help in identifying influential nodes in a social media network. The three key centrality measures used in this project are:

#### **1. Degree Centrality**

Degree centrality measures the number of direct connections a node has. It helps identify highly connected users in the network.

* **Formula:**CD(v)=deg(v)N−1C\_D(v) = \frac{\text{deg}(v)}{N - 1}CD​(v)=N−1deg(v)​  
  where:
  + CD(v)C\_D(v)CD​(v) is the degree centrality of node vvv.
  + deg(v)\text{deg}(v)deg(v) is the number of edges connected to node vvv.
  + NNN is the total number of nodes in the network.
* **Interpretation:**
  + A higher degree centrality means the node has many direct connections.
  + Influencers on platforms like Twitter or Instagram often have high degree centrality.

#### **2. Closeness Centrality**

Closeness centrality measures how quickly a node can reach other nodes in the network. It is based on the shortest path between nodes.

* **Formula:**CC(v)=N−1∑u≠vd(v,u)C\_C(v) = \frac{N - 1}{\sum\_{u \neq v} d(v, u)}CC​(v)=∑u=v​d(v,u)N−1​  
  where:
  + CC(v)C\_C(v)CC​(v) is the closeness centrality of node vvv.
  + d(v,u)d(v, u)d(v,u) is the shortest distance between nodes vvv and uuu.
  + NNN is the total number of nodes in the network.
* **Interpretation:**
  + A node with high closeness centrality can spread information quickly.
  + Useful for identifying efficient information disseminators in social networks.

#### **3. Betweenness Centrality**

Betweenness centrality measures how often a node lies on the shortest path between other nodes. It identifies nodes that act as bridges in the network.

* **Formula:**CB(v)=∑s≠v≠tσst(v)σstC\_B(v) = \sum\_{s \neq v \neq t} \frac{\sigma\_{st}(v)}{\sigma\_{st}}CB​(v)=s=v=t∑​σst​σst​(v)​  
  where:
  + CB(v)C\_B(v)CB​(v) is the betweenness centrality of node vvv.
  + σst\sigma\_{st}σst​ is the number of shortest paths from node sss to node ttt.
  + σst(v)\sigma\_{st}(v)σst​(v) is the number of those paths that pass through node vvv.
* **Interpretation:**
  + Nodes with high betweenness centrality control information flow.
  + Important for identifying gatekeepers in social networks.

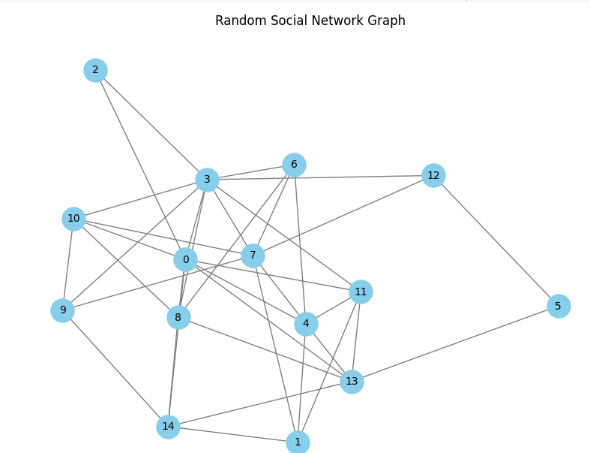
### **Identifying the Most Influential Node**

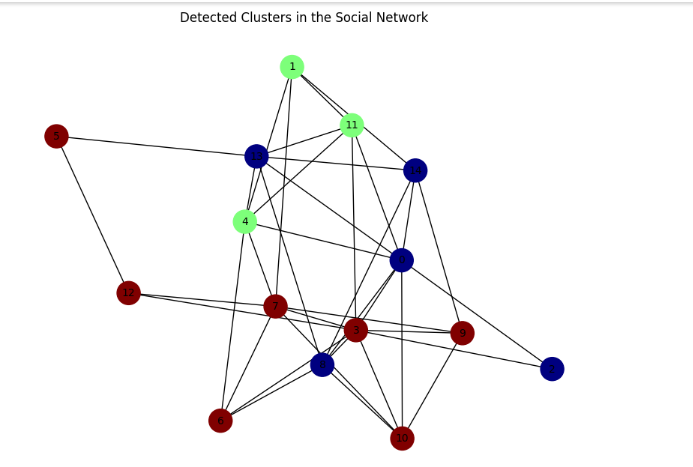
The most influential node in a social media network can be determined based on:

1. **Highest Degree Centrality:** Indicates direct influence due to a high number of connections.
2. **Highest Closeness Centrality:** Indicates the ability to quickly reach all other nodes.
3. **Highest Betweenness Centrality:** Indicates control over information flow.

A node that scores highly in all three metrics is likely the most influential.

**OUTPUT:**

[**SMA\_exp6.ipynb**](https://colab.research.google.com/drive/1DlNTZ1I_EGkVtSlVWTi9uJlpqbzroofO?authuser=0#scrollTo=sXBY_VeQeeBQ) ****

****

**Conclusion:**

The experiment successfully demonstrated the use of **NetworkX** for social network analysis and centrality computation. By applying these centrality metrics, we identified key influencers within the network. This approach is highly useful in real-world applications such as **viral marketing, influencer identification, social media monitoring, and opinion leader detection.** The results highlight how structural properties of a network can provide deep insights into the role of individuals in social media interactions.