Link Prediction Algorithms for Multilayer Networks

Harsh Yadav (20)

Supervisor: Dr. Kuldeep Singh

Department of Computer Science
Faculty of Mathematical Sciences
University of Delhi

November 2024

Overview

1. Introduction to Link prediction

What is Link Prediction

- Link prediction is a task in network analysis that aims at predicting the likelihood of future connections between nodes in a network **[1]**.
- It uses the existing network structure and sometimes node attributes to estimate the presence or absence of potential links.
- This task is often applied to graphs G(V, E), where V represents the set of nodes as entities and E represents the set of edges as connections or links.
- The network's existing relationships help predict potential connections between unconnected nodes.

What is Link Prediction

- Link prediction is a task in network analysis that aims at predicting the likelihood of future connections between nodes in a network **[1]**.
- It uses the existing network structure and sometimes node attributes to estimate the presence or absence of potential links.
- This task is often applied to graphs G(V, E), where V represents the set of nodes as entities and E represents the set of edges as connections or links.
- The network's existing relationships help predict potential connections between unconnected nodes.

What is Link Prediction

- Link prediction is a task in network analysis that aims at predicting the likelihood of future connections between nodes in a network **[1]**.
- It uses the existing network structure and sometimes node attributes to estimate the presence or absence of potential links.
- This task is often applied to graphs G(V, E), where V represents the set of nodes as entities and E represents the set of edges as connections or links.
- The network's existing relationships help predict potential connections between unconnected nodes.

What is Link Prediction

- Link prediction is a task in network analysis that aims at predicting the likelihood of future connections between nodes in a network **[1]**.
- It uses the existing network structure and sometimes node attributes to estimate the presence or absence of potential links.
- This task is often applied to graphs G(V, E), where V represents the set of nodes as entities and E represents the set of edges as connections or links.
- The network's existing relationships help predict potential connections between unconnected nodes.

Example of Link Prediction

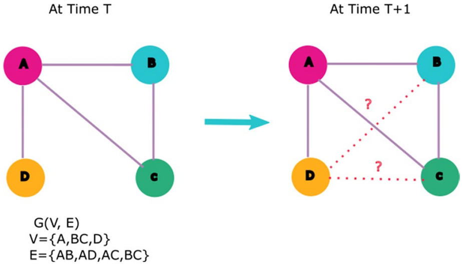


Fig. 1 Example of link prediction in network graph

Consider the example in Fig. 1. If node A is linked to both nodes B and D, nodes B and D are not linked, and node A is a common link between B and D, then link prediction would suggest a potential connection between D and B, as well as C and D.

Networks used in link prediction

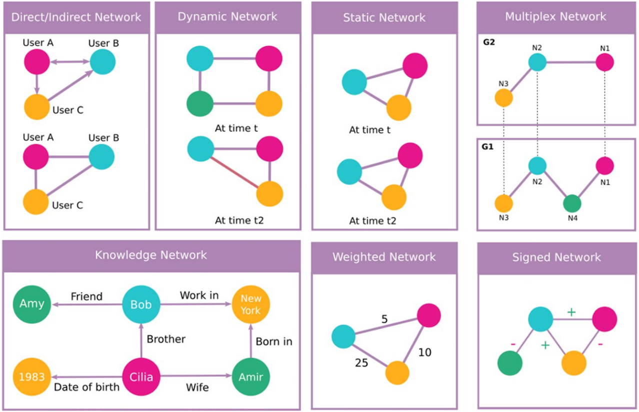


Fig. 2 Representation of different types of networks

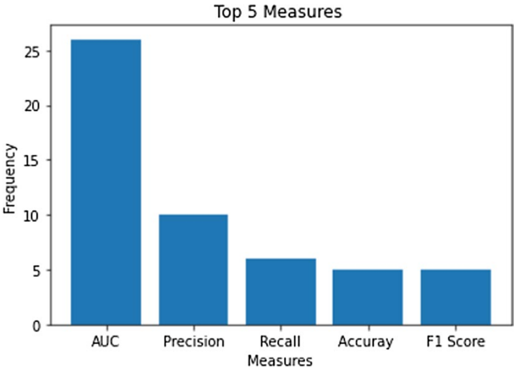
The most prevalent networks employed in link prediction approaches are represented in Fig. 2.

Popular datasets used in link prediction

Table 4 Description of popular datasets used in link prediction

Dataset	Nodes	Edges	Dir/Undir	Type of networks
Karate	34	78	Undirected	Static Network
Dolphin	62	159	Undirected	Static Network
Jazz	198	2742	Undirected	Static
Football	242	4100	Directed	Directed Network
USAir	332	2100	Undirected	Static Network
Email-Eu-core	1005	25,571	Directed	Dynamic Network
CollegeMsg	1899	T59835, S20296	Undirected	Dynamic Network
Yeast	2375	11,693	Undirected	Static Network
Facebook NIPS	2888	2981	Undirected	Ego Network
The Ego-Facebook	4039	88,234	Undirected	Ego Network
Power	4941	6594	Directed	Static Network
Twitter	1043	4860	Directed	Static Network
Twitter-Foursquare	10,989	16,104, 23,348, 11,459	Directed	Multiplex network
Vickers	29	240, 126, 152	Undirected	Multiplex

Evaluation metrics



Top Universities interested in Link Prediction research

Table 2 Top Universities in countries of interest for link prediction research

Country	University
USA	Stanford University
China	University of Electronic Science and Technology of China
India	The Department of Science and Technology (DST)
USA	Carnegie Mellon University
France	CNRS Centre National de la Recherche Scientifique
Germany	Max Planck Institute for Informatics

2. Algorithm-1

Common Neighbor-based Link Prediction (CN)

- The Common Neighbor approach was adapted for multilayer networks by Berlingerio et al. in their 2011 paper **[2]**.
- For two nodes u and v, the likelihood of a link between them is given by:

$$CN(u,v) = \sum_{l} |\Gamma_l(u) \cap \Gamma_l(v)|$$

- In a multilayer network, where $\Gamma_l(u)$ and $\Gamma_l(v)$ are the sets of neighbors for nodes u and v in layer l.
- Consider a multilayer network with two layers L_1 and L_2:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Common Neighbor-based Link Prediction (CN)

- The Common Neighbor approach was adapted for multilayer networks by Berlingerio et al. in their 2011 paper**[2]**.
- For two nodes u and v, the likelihood of a link between them is given by:

$$CN(u, v) = \sum_{l} |\Gamma_l(u) \cap \Gamma_l(v)|$$

- In a multilayer network, where $\Gamma_l(u)$ and $\Gamma_l(v)$ are the sets of neighbors for nodes u and v in layer l.
- Consider a multilayer network with two layers L_1 and L_2:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Common Neighbor-based Link Prediction (CN)

- The Common Neighbor approach was adapted for multilayer networks by Berlingerio et al. in their 2011 paper**[2]**.
- For two nodes u and v, the likelihood of a link between them is given by:

$$CN(u,v) = \sum_{l} |\Gamma_l(u) \cap \Gamma_l(v)|$$

- In a multilayer network, where $\Gamma_l(u)$ and $\Gamma_l(v)$ are the sets of neighbors for nodes u and v in layer l.
- Consider a multilayer network with two layers L_1 and L_2:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Common Neighbor-based Link Prediction (CN)

- The Common Neighbor approach was adapted for multilayer networks by Berlingerio et al. in their 2011 paper**[2]**.
- For two nodes u and v, the likelihood of a link between them is given by:

$$CN(u,v) = \sum_{l} |\Gamma_l(u) \cap \Gamma_l(v)|$$

- In a multilayer network, where $\Gamma_l(u)$ and $\Gamma_l(v)$ are the sets of neighbors for nodes u and v in layer l.
- Consider a multilayer network with two layers L_1 and L_2:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Common Neighbor-based Link Prediction (CN)

- The Common Neighbor approach was adapted for multilayer networks by Berlingerio et al. in their 2011 paper**[2]**.
- For two nodes u and v, the likelihood of a link between them is given by:

$$CN(u,v) = \sum_{l} |\Gamma_l(u) \cap \Gamma_l(v)|$$

- In a multilayer network, where $\Gamma_l(u)$ and $\Gamma_l(v)$ are the sets of neighbors for nodes u and v in layer l.
- Consider a multilayer network with two layers L_1 and L_2:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Common Neighbor-based Link Prediction (CN)

- The Common Neighbor approach was adapted for multilayer networks by Berlingerio et al. in their 2011 paper**[2]**.
- For two nodes u and v, the likelihood of a link between them is given by:

$$CN(u, v) = \sum_{l} |\Gamma_l(u) \cap \Gamma_l(v)|$$

- In a multilayer network, where $\Gamma_l(u)$ and $\Gamma_l(v)$ are the sets of neighbors for nodes u and v in layer l.
- Consider a multilayer network with two layers L_1 and L_2:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Remarks

- To predict a link between nodes A and C:

$$CN(A, C) = |\Gamma_{L_1}(A) \cap \Gamma_{L_1}(C)| + |\Gamma_{L_2}(A) \cap \Gamma_{L_2}(C)| = 1 + 0 = 1$$

- This indicates a potential link between nodes A and C.

Remarks

- To predict a link between nodes A and C:

$$CN(A, C) = |\Gamma_{L_1}(A) \cap \Gamma_{L_1}(C)| + |\Gamma_{L_2}(A) \cap \Gamma_{L_2}(C)| = 1 + 0 = 1$$

- This indicates a potential link between nodes A and C.

3. Algorithm-2

Katz Index-based Link Prediction (Katz)

- The Katz similarity was introduced by Katz in 1953 and has since been adapted for multilayer networks.
- The score for node pair u and v is given by:

$$\mathrm{Katz}(u,v) = \sum_{p=1}^{\infty} \beta^p (\mathcal{A}^p)_{uv}$$

where $0 < \beta < 1$ is a damping factor.
- In multilayer networks, construct the supra-adjacency matrix $\mathcal{A}$.
- For example, consider two layers represented by adjacency matrices:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Katz Index-based Link Prediction (Katz)

- The Katz similarity was introduced by Katz in 1953 and has since been adapted for multilayer networks.
- The score for node pair u and v is given by:

$$\text{Katz}(u, v) = \sum_{p=1}^{\infty} \beta^p (\mathcal{A}^p)_{uv}$$

where $0 < \beta < 1$ is a damping factor.

- In multilayer networks, construct the supra-adjacency matrix $\mathcal{A}$.
- For example, consider two layers represented by adjacency matrices:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Katz Index-based Link Prediction (Katz)

- The Katz similarity was introduced by Katz in 1953 and has since been adapted for multilayer networks.
- The score for node pair u and v is given by:

$$\mathrm{Katz}(u, v) = \sum_{p=1}^{\infty} \beta^{p} (\mathcal{A}^{p})_{uv}$$

  where $0 < \beta < 1$ is a damping factor.
- In multilayer networks, construct the supra-adjacency matrix $\mathcal{A}$.
- For example, consider two layers represented by adjacency matrices:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Katz Index-based Link Prediction (Katz)

- The Katz similarity was introduced by Katz in 1953 and has since been adapted for multilayer networks.
- The score for node pair u and v is given by:

$$\text{Katz}(u, v) = \sum_{p=1}^{\infty} \beta^p (\mathcal{A}^p)_{uv}$$

where $0 < \beta < 1$ is a damping factor.

- In multilayer networks, construct the supra-adjacency matrix $\mathcal{A}$.
- For example, consider two layers represented by adjacency matrices:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Katz Index-based Link Prediction (Katz)

- The Katz similarity was introduced by Katz in 1953 and has since been adapted for multilayer networks.
- The score for node pair u and v is given by:

$$\mathrm{Katz}(u, v) = \sum_{p=1}^{\infty} \beta^p (\mathcal{A}^p)_{uv}$$

where $0 < \beta < 1$ is a damping factor.

- In multilayer networks, construct the supra-adjacency matrix $\mathcal{A}$.
- For example, consider two layers represented by adjacency matrices:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Katz Index-based Link Prediction (Katz)

- The Katz similarity was introduced by Katz in 1953 and has since been adapted for multilayer networks.
- The score for node pair u and v is given by:

$$\text{Katz}(u, v) = \sum_{p=1}^{\infty} \beta^p (\mathcal{A}^p)_{uv}$$

  where $0 < \beta < 1$ is a damping factor.
- In multilayer networks, construct the supra-adjacency matrix $\mathcal{A}$.
- For example, consider two layers represented by adjacency matrices:
- Layer L_1:

$$A^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Layer L_2:

$$A^{(2)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Katz Score Calculation

- The supra-adjacency matrix $\mathcal{A}$ is:

$$\mathcal{A} = A^{(1)} + A^{(2)} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

- Let's choose $\beta = 0.1$.
- Now we calculate the Katz score for nodes A and C.
- **First Power** ($p = 1$):

$$\mathcal{A}^1 = \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 1) = 0.1 \cdot (\mathcal{A}^1)_{AC} = 0.1 \cdot 1 = 0.1$$

- **Second Power** ($p = 2$): Calculate $\mathcal{A}^2$:

$$\mathcal{A}^2 = \mathcal{A} \times \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 2) = 0.1^2 \cdot (\mathcal{A}^2)_{AC} = 0.01 \cdot 3 = 0.03$$

Katz Score Calculation

- The supra-adjacency matrix $\mathcal{A}$ is:

$$\mathcal{A} = A^{(1)} + A^{(2)} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

- Let's choose $\beta = 0.1$.
- Now we calculate the Katz score for nodes A and C.
- **First Power ($p = 1$)**:

$$\mathcal{A}^1 = \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 1) = 0.1 \cdot (\mathcal{A}^1)_{AC} = 0.1 \cdot 1 = 0.1$$

- **Second Power ($p = 2$)**: Calculate $\mathcal{A}^2$:

$$\mathcal{A}^2 = \mathcal{A} \times \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 2) = 0.1^2 \cdot (\mathcal{A}^2)_{AC} = 0.01 \cdot 3 = 0.03$$

Katz Score Calculation

- The supra-adjacency matrix $\mathcal{A}$ is:

$$\mathcal{A} = A^{(1)} + A^{(2)} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

- Let's choose $\beta = 0.1$.
- Now we calculate the Katz score for nodes A and C.
- **First Power ($p = 1$):**

$$\mathcal{A}^1 = \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 1) = 0.1 \cdot (\mathcal{A}^1)_{AC} = 0.1 \cdot 1 = 0.1$$

- **Second Power ($p = 2$):** Calculate $\mathcal{A}^2$:

$$\mathcal{A}^2 = \mathcal{A} \times \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 2) = 0.1^2 \cdot (\mathcal{A}^2)_{AC} = 0.01 \cdot 3 = 0.03$$

Katz Score Calculation

- The supra-adjacency matrix $\mathcal{A}$ is:

$$\mathcal{A} = A^{(1)} + A^{(2)} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

- Let's choose $\beta = 0.1$.
- Now we calculate the Katz score for nodes A and C.
- **First Power ($p = 1$)**:

$$\mathcal{A}^1 = \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 1) = 0.1 \cdot (\mathcal{A}^1)_{AC} = 0.1 \cdot 1 = 0.1$$

- **Second Power ($p = 2$)**: Calculate $\mathcal{A}^2$:

$$\mathcal{A}^2 = \mathcal{A} \times \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 2) = 0.1^2 \cdot (\mathcal{A}^2)_{AC} = 0.01 \cdot 3 = 0.03$$

Katz Score Calculation

- The supra-adjacency matrix $\mathcal{A}$ is:

$$\mathcal{A} = A^{(1)} + A^{(2)} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

- Let's choose $\beta = 0.1$.
- Now we calculate the Katz score for nodes A and C.
- **First Power ($p = 1$)**:

$$\mathcal{A}^1 = \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 1) = 0.1 \cdot (\mathcal{A}^1)_{AC} = 0.1 \cdot 1 = 0.1$$

- **Second Power ($p = 2$)**: Calculate $\mathcal{A}^2$:

$$\mathcal{A}^2 = \mathcal{A} \times \mathcal{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 2) = 0.1^2 \cdot (\mathcal{A}^2)_{AC} = 0.01 \cdot 3 = 0.03$$

Katz Score Calculation

- **Third Power ($p = 3$)**: Calculate $\mathcal{A}^3$:

$$\mathcal{A}^3 = \mathcal{A}^2 \times \mathcal{A} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 14 & 10 \\ 10 & 10 & 7 \\ 7 & 10 & 5 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 3) = 0.1^3 \cdot (\mathcal{A}^3)_{AC} = 0.001 \cdot 10 = 0.01$$

- In practice, we continue calculating until the contributions become negligible. The contributions from the first few powers, as calculated above, already provide a reasonable estimate of the Katz score.
- **Result:**
Thus, the Katz score Katz(A, C) for nodes A and C is a combination of these contributions:

$$\text{Katz}(A, C) = 0.1 + 0.03 + 0.01 + \dots$$

This indicates the potential for a link between A and C based on the number and length of paths connecting them through the network.

Katz Score Calculation

- **Third Power ($p = 3$)**: Calculate $\mathcal{A}^3$:

$$\mathcal{A}^3 = \mathcal{A}^2 \times \mathcal{A} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 14 & 10 \\ 10 & 10 & 7 \\ 7 & 10 & 5 \end{pmatrix}$$

  For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 3) = 0.1^3 \cdot (\mathcal{A}^3)_{AC} = 0.001 \cdot 10 = 0.01$$

- In practice, we continue calculating until the contributions become negligible. The contributions from the first few powers, as calculated above, already provide a reasonable estimate of the Katz score.
- **Result:**
  Thus, the Katz score $\text{Katz}(A, C)$ for nodes A and C is a combination of these contributions:

$$\text{Katz}(A, C) = 0.1 + 0.03 + 0.01 + \dots$$

  This indicates the potential for a link between A and C based on the number and length of paths connecting them through the network.

Katz Score Calculation

- **Third Power ($p = 3$)**: Calculate $\mathcal{A}^3$:

$$\mathcal{A}^3 = \mathcal{A}^2 \times \mathcal{A} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 14 & 10 \\ 10 & 10 & 7 \\ 7 & 10 & 5 \end{pmatrix}$$

For A to C:

$$\text{Katz}(A, C) \text{ (contribution from } p = 3) = 0.1^3 \cdot (\mathcal{A}^3)_{AC} = 0.001 \cdot 10 = 0.01$$

- In practice, we continue calculating until the contributions become negligible. The contributions from the first few powers, as calculated above, already provide a reasonable estimate of the Katz score.
- **Result:**
  Thus, the Katz score $\text{Katz}(A, C)$ for nodes A and C is a combination of these contributions:

$$\text{Katz}(A, C) = 0.1 + 0.03 + 0.01 + \dots$$

This indicates the potential for a link between A and C based on the number and length of paths connecting them through the network.

References

- **[1]** Arrar, D., Kamel, N., & Lakhfif, A. (2024). A comprehensive survey of link prediction methods. The Journal of Supercomputing, 80(3), 3902-3942. https://doi.org/10.1007/s11227-023-05591-8
- **[2]** Berlingerio, Michele & Coscia, Michele & Giannotti, Fosca & Monreale, Anna & Pedreschi, Dino. (2011). Foundations of Multidimensional Network Analysis. Proceedings - 2011 International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2011. 485-489. 10.1109/ASONAM.2011.103.
- **[3]** Mojarad, Musa Dehghan, Omid Dehghan, Amin Salar, Alireza. (2021). Development of Katz Similarity Index to Improve Link Prediction in Social Networks Based on Reliable Routes. 9. 16-20.

Thank You