Network Analysis Methods 2

Network Science

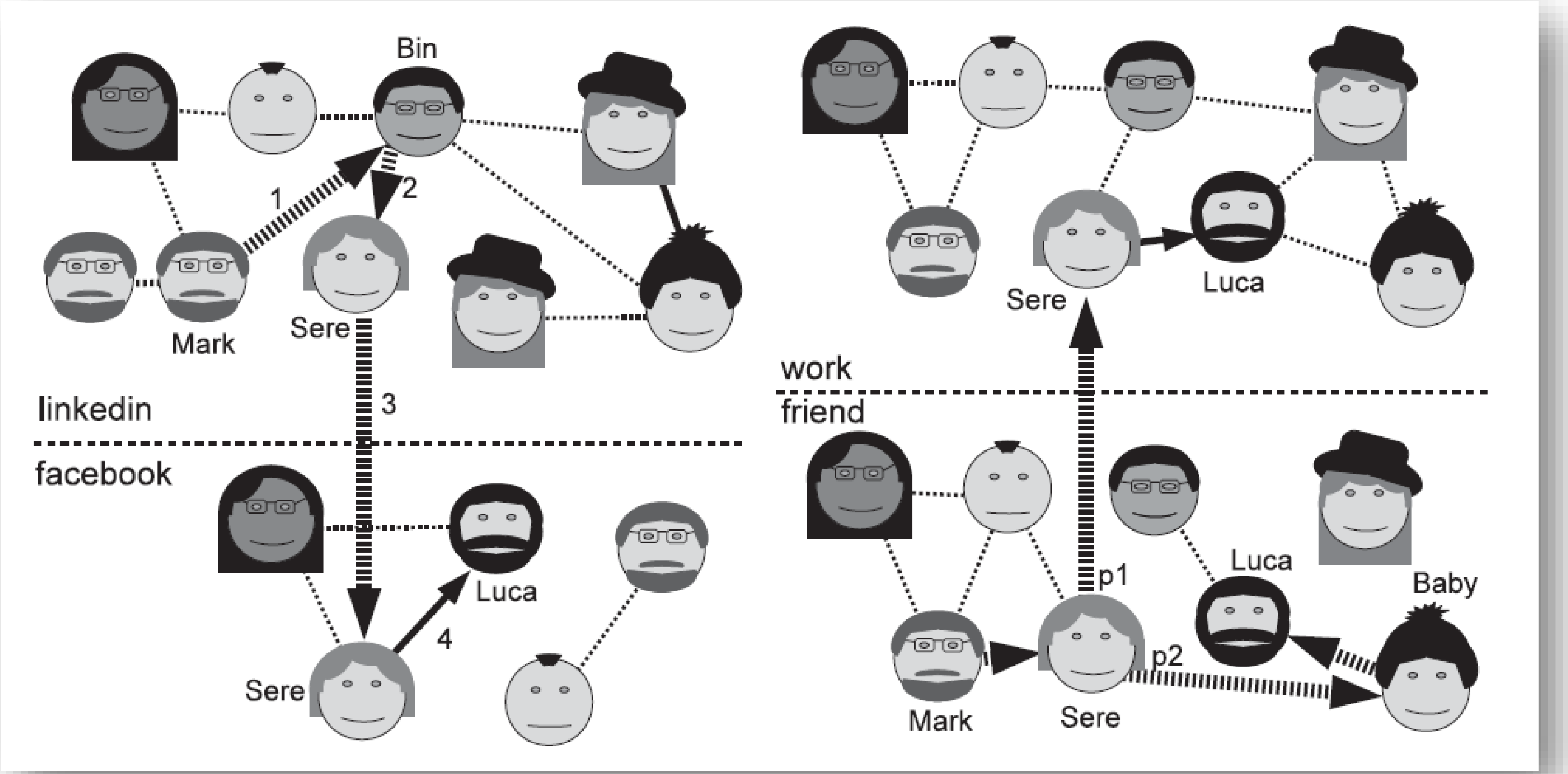
Multilayer Social Networks
Distance-based Measures and Flattening
2023/24

Outline

- Distance based measures
- Measures of relevance
- Flattening and projection

Random Walk

- In the context of multilayer network, random walks are processes describing an item – often called a walker – randomly flowing through the available relational ties and also able to switch layers.



Degree Centrality (reminder)

- **Degree Centrality:** Let $a \in A$ be an actor an $L \subseteq \boldsymbol{L}$ a set of layers and $M = (A, \mathbf{L}, V, E)$ a multilayer network. The degree centrality of a on L is defined as

$$degree(a, L) = |\{\{(a,l),(a',l')\} \in E \ s.t. \ l, l' \in L\}|.$$

Random Walk-Based Extensions of Degree Centrality

- Random walk-based methods offer a powerful tool to define multiple measures on top of a single basic concept.

- Within this approach, a measure related to degree centrality in multilayer networks is *occupation centrality*.

- **Occupation centrality:** The *occupation centrality* of an actor $a \in A$ in a multilayer network $M = (A, \mathbf{L}, V, E)$ is the probability that a random walker on M is found on any node corresponding to a.

Random Walk-Based Extensions of Degree Centrality

- In this case, a walker in node *(a, I)* might jump to one of its neighbors *(a′, I)* within the same layer *l* with uniform probability or might switch to its counterpart *(a , I‘)* on a different layer *l‘*.

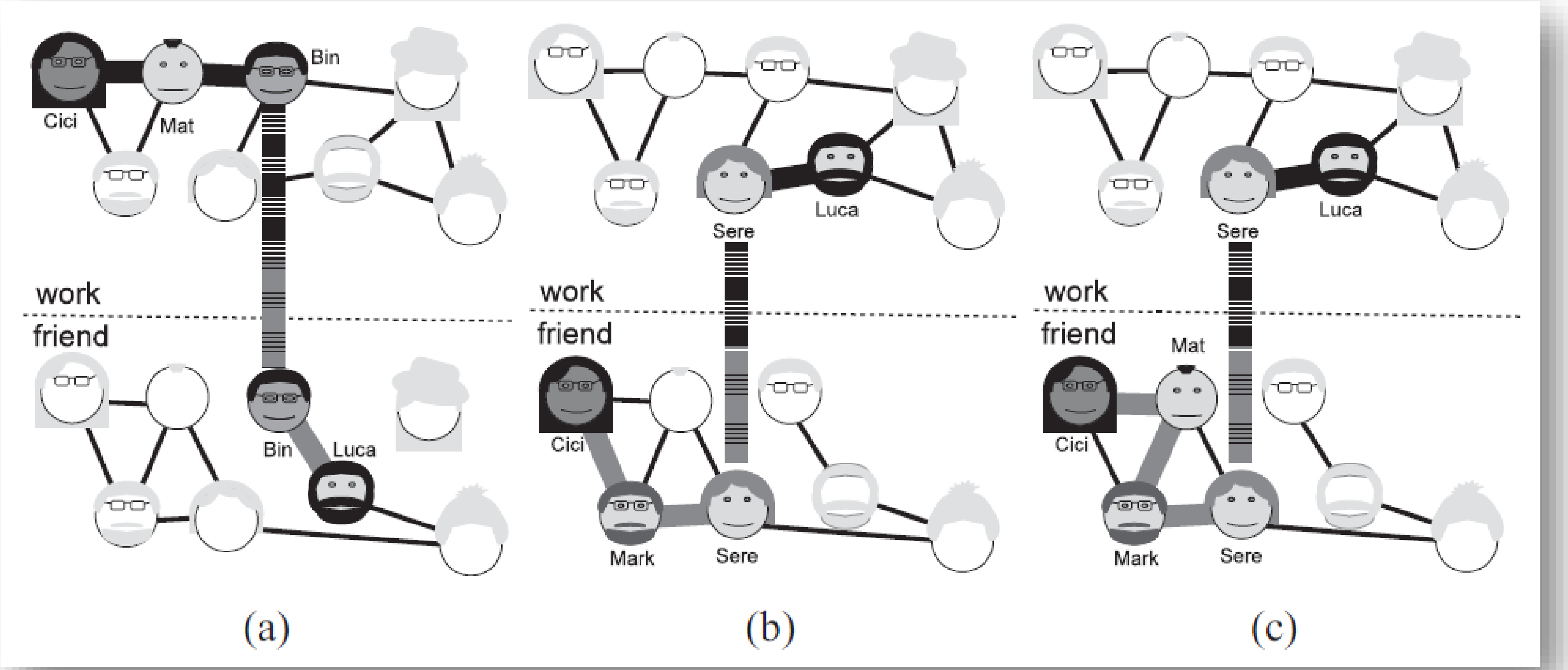
- This formulation can be easily extended for weighted networks assuming a layer-switching probability proportional to the strength of the connection.

- In the specific case when interlayer edges have the same strength for all nodes, occupation centrality is strongly correlated with degree centrality.

Distance Based Measures

- A more general concept of *Multilayer Distance* that makes a difference between edges on different layers.
- The distance between two actors is one of the most complex and interesting concepts we need to extend to define new measures on multilayer networks.
- A first consideration is that the shortest path might not be the best option for evaluating the separation between the two actors.
- Under the availability of multiple layers, instead of using a shortest path, actors could prefer to switch layers and reach their destination through a longer, but maybe more reliable, path.

Three alternative paths from Cici to Luca



- (a) and (b) have the same number of steps but different lengths on different layers, whereas (c) is longer than (b).

Paths and Path Lengths

- **Paths from Cici to Luca:** The first consists of two steps on the *work* layer, a layer switch from *work* to *friend*, and one step on the *friend* layer, whereas the second has two friend steps, a layer switch from *friend* to *work*, and one *work* step.

- **Multilayer path length:** The multilayer length of a *p* on layers $\boldsymbol{L} = \{l_1, \ldots, l_m\}$ is a matrix L where L_{ij} indicates the number of edges traversed from a node in layer l_i to a node in layer l_j.

Cici to Luca…

- The lengths of the two paths in (a) and (b) can be represented as

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$$

- where the first row/column of each matrix represents the *work* layer and the second row/column represents the *friend* layer.

Different representations

- We can obtain different representations of the lengths, depending on the level of detail we want to keep.
- For example, we can omit interlayer switching costs, only keeping the diagonal of the length matrix.
- In this case, our two distances can be expressed using the following vectors:

$$\begin{pmatrix} 2 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \end{pmatrix}$$

- As an extreme solution, we can sum all values into a single number, reducing our multilayer concept to a traditional length – in our example, both paths would appear having the same length.

Shorter-than…

- While paths (a) and (b) are incomparable, path (b) is shorter than path (c), because it involves a shorter path on the friend layer and the same number of steps on the work layer.

- **Shorter-than relation:** Let *R* and *S* be two multilayer path lengths. *R* is shorter than *S* if and only if

$$\forall i, j \; R_{ij} \leq S_{ij} \wedge \exists i j \; R_{ij} < S_{ij}.$$

Multilayer Shortest Paths

- Multilayer shortest paths maintain on a general level the idea that a shorter path is on average better than a longer path.
  - The distance between two actors is expressed as a set of incomparable shortest paths that can traverse multiple layers.
  - This concept is easily extensible to allow weighted edges by replacing the number of steps with the sum of weights along the path.

Closeness and Betweenness Centrality in Multilayer Networks

- **Random walk closeness:** The *closeness centrality* of an actor $a \in A$ is defined as the inverse of the average number of steps that a random walker, starting from any other actor in the multilayer network, requires to hit a for the first time.

- **Random walk betweenness:** Given an actor $a \in A$, the *betweenness centrality* is defined as the number of random walks between any pair of nodes that pass through any node corresponding to a, averaging the value over all possible starting layers.

Measures of Relevance

- One of the most intriguing aspects of multilayer network analysis is its ability to understand the relation between an actor and a specific layer as well as between different layers within the same multilayer network.

- If we want to consider multiple layers, we must be able to explore how single layers relate to the whole network structure and to what extent a single layer is an important part of an actor's social network.

Neighborhood Based Measures (reminder)

- **Neighbors:** Let $a \in A$ be an actor an $L \subseteq \boldsymbol{L}$ a set of layers and $M = (A, \boldsymbol{L}, V, E)$ a multilayer network. The neighbors of a on layers L are defined as

$$neighbors(a, L) = \{a' \in A | \{(a, l), (a', l')\} \in E \text{ } and \text{ } l, l' \in L\}.$$

- **Neighborhood Centrality:** The neighborhood of a on layers L is defined as

$$neighborhood(a, L) = |neighbors(a, L)|.$$

- **Connective Redundancy:** The connective redundancy of a on layers L is defined as

$$connective\ redundancy(a, L) = 1 - \frac{neighborhood(a, L)}{degree(a, L)}.$$

- **Exclusive Neighborhood:** The exclusive neighborhood of a on layers L is defined as

$$xneighborhood(a, L) = |neighbors(a, L) \setminus neighbors(a, \boldsymbol{L} \setminus L)|.$$

Relevance

- **Relevance:** Let $a \in A$ be an actor, $L \subseteq \boldsymbol{L}$ a set of layers and $M = (A, \boldsymbol{L}, V, E)$ a multilayer network. *Relevance* is defined as follows:

$$\text{relevance}(a, L) = \frac{\text{neighborhood}(a, L)}{\text{neighborhood}(a, \mathcal{L})}$$

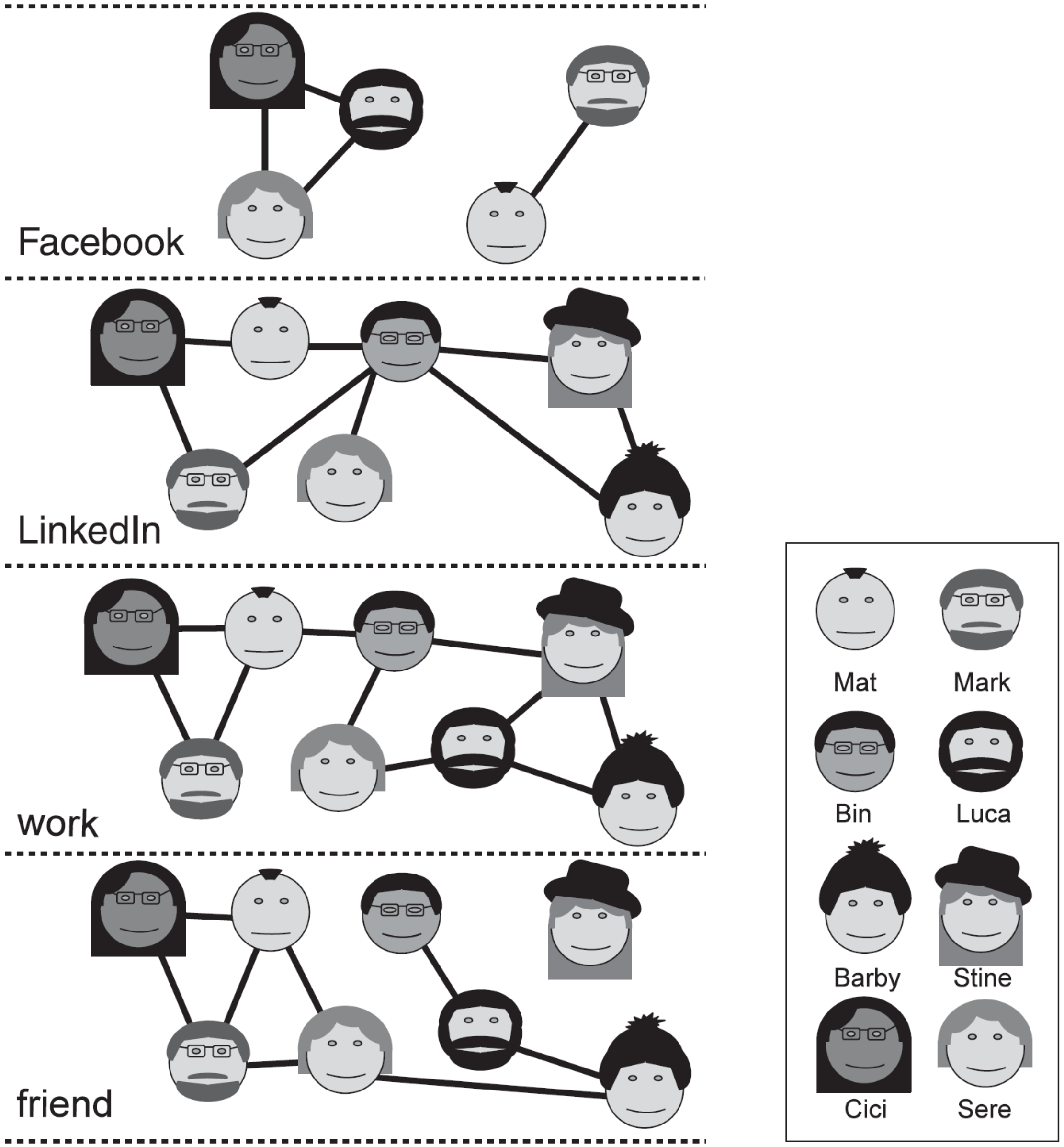
  - *Relevance* computes the ratio between the neighbors of an actor connected by edges belonging to a specific set of layers and the total number of her neighbors.
  - The set L might also contain only a single layer, of which we might want to study the specific role within the multilayer network.

Exclusive Version of Relevance

- Exclusive layer relevance: Let $a \in A$ be an actor, $L \subseteq \boldsymbol{L}$ a set of layers and $M = (A, \boldsymbol{L}, V, E)$ a multilayer network. *Exclusive relevance* is defined as follows:

$$\text{xrelevance}(a, L) = \frac{\text{xneighborhood}(a, L)}{\text{neighborhood}(a, \mathcal{L})}$$

  - *Exclusive relevance* computes the fraction of neighbors directly connected with actor a through edges belonging only to layers in *L*.


Facebook
LinkedIn
work
friend
Mat
Mark
Bin
Luca
Barby
Stine
Cici
Sere

Relevance for the actors in the example

	User	Facebook	Friend	LinkedIn	Work
1	Cici	0.50	0.50	0.50	0.50
2	Mat	0.25	0.75	0.50	0.75
3	Mark	0.25	0.75	0.50	0.50
4	Bin		0.17	0.83	0.50
5	Serena	0.33	0.50	0.17	0.33
6	Barby		0.50	0.50	0.50
7	Stine			0.67	1.00
8	Luca	0.40	0.40		0.60

Exclusive relevance for the actors

	User	Facebook	Friend	LinkedIn	Work
1	Cici	0.50	0.00	0.00	0.00
2	Mat	0.00	0.25	0.00	0.00
3	Mark	0.00	0.25	0.25	0.00
4	Bin		0.17	0.33	0.00
5	Serena	0.17	0.50	0.00	0.00
6	Barby		0.25	0.25	0.00
7	Stine			0.00	0.33
8	Luca	0.20	0.20		0.20

Flattening and Projection

- A basic approach to dealing with multilayer networks is to reconstruct a (weighted) single-layer social network so that existing methods, such as community detection, can be directly applied.
  - If nodes on the different layers correspond to a common set of actors, we normally talk of flattening, that is, the process of merging all nodes corresponding to the same actor into a single node.
  - When we have multiple types of nodes, a common operation consists in projecting the network only on one type of node, discarding the others.

Flattening

- A basic flattening process consists in creating a layer with one node for each actor and an edge between two nodes *if an edge among two nodes corresponding to those actors exists somewhere in the multilayer network*.

- **Basic flattening:** A *basic* (unweighted) *flattening* of a multilayer network $M = (A, \mathbf{L}, V, E)$ is a graph (V_f, E_f), where $V_f = \{a \mid (a, l) \in V\}$ and $E_f = \{(a_i, a_j) \mid \{(a_i, l_q), (a_j, l_r)\} \in E\}$.

Weighted Version of Flattening

- A simple variation of basic flattening consists in adding a weight to each edge in the flattened network proportional to the number or edges between the actors corresponding to those nodes.

- A more general approach consists in assigning a weight Θ_{qr} to each pair of layers (l_q, l_r), so that the resulting single-layer network can be expressed as a linear combination of the original multilayer network.

Weighted Flattening

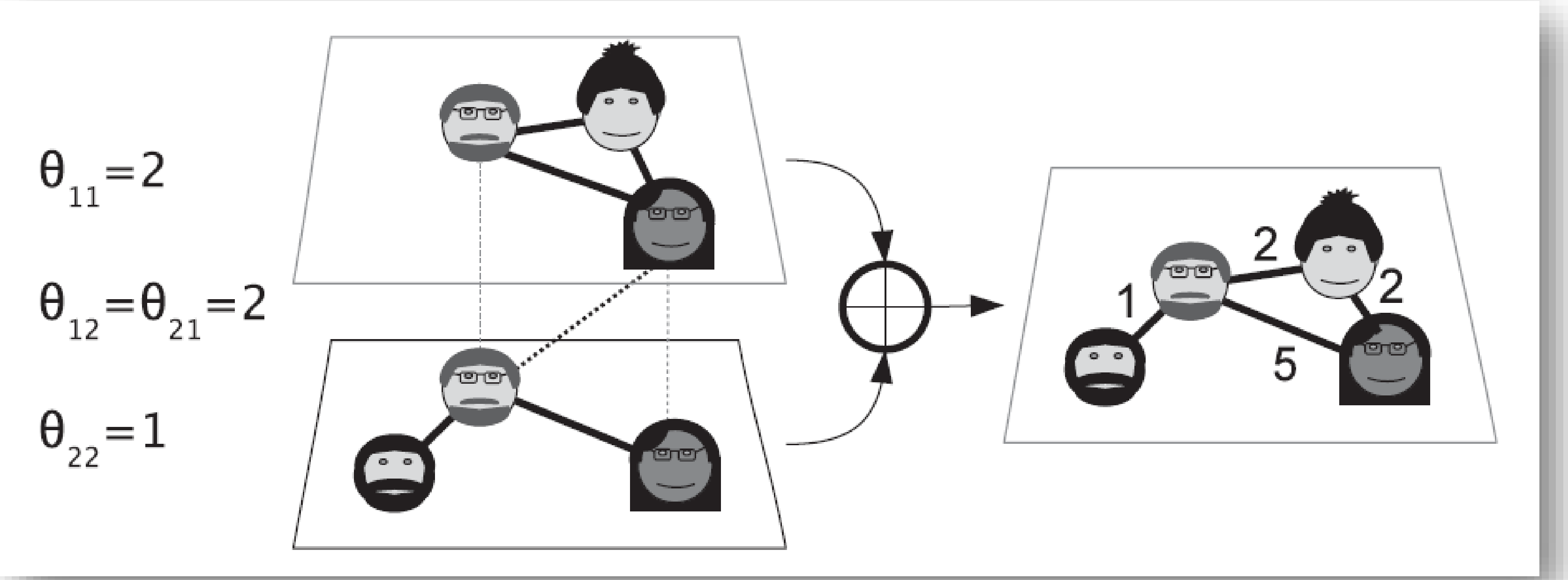
- Weighted flattening: Let $M = (A, \mathbf{L}, V, E)$ be a multilayer network. Given a $|\mathbf{L}| \; x \; |\mathbf{L}|$ matrix Θ, where Θ_{qr} indicates the weight to be assigned to edges from layer l_q to layer l_r, a weighted flattening of M is a weighted graph (V_f, E_f, ω), where (V_f, E_f) is a basic flattening of M and

$$\omega(a_i, a_j) = \sum_{\{(a_i,l_q),(a_j,l_r)\}\in E} \theta_{q,r}.$$

- This definition generalizes the simple weighted flattening strategy, which can be expressed by setting $\Theta_{qr} = 1$ and can also be used to remove some of the layers by setting their weights to 0.

Example

- Weighted flattening of two layers, with varying weights depending on the starting and ending layers of edges in the original multilayer network.



Projection

- Network projection is a traditional approach used to simplify two-mode networks, which can be modeled as multilayer networks where each node type (say, A or B) is represented by a layer.

- The most straightforward approach to network projection consists in removing type B nodes and adding an edge between any pair of type A nodes originally connected to the same type B node.

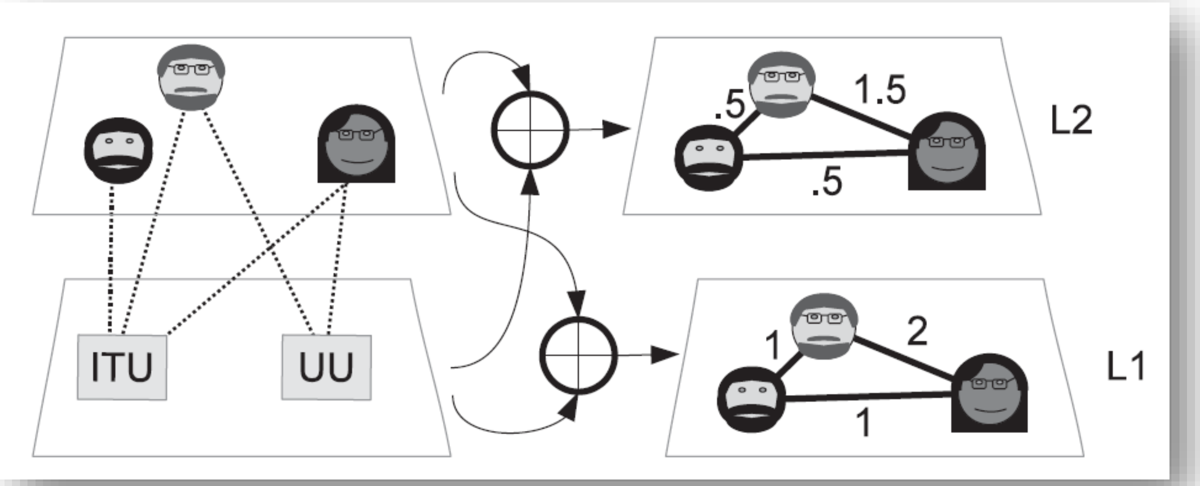
Example

- **L1:** Type A nodes get connected by a weighted edge with weight w defined as

$$w(i, j) = \sum_p 1,$$

- where p indicates the type B nodes connected to both i and j.

- L2: A different weight assignment based on $w(i, j) = \sum_p \frac{1}{N_p - 1}$, where N_p is the number of nodes of type A connected through the p-th type B node.



Seminar Assignment

- Select one multi-layer network from the previously listed sources and compute, for each network actor/node, relevance measures given in the presentation.

- Implement random walk processes in the selected multi-layer network and, based on it, compute occupation centrality for each network node.

- Implement unweighted flattening and apply it to the selected network.

References

- Dickison, M. E., Magnani, M., Rossi, L. (2016). *Multilayer social networks*. Cambridge University Press. http://multilayer.it.uu.se.

- Bianconi, G. (2018). *Multilayer networks: structure and function*. Oxford university press.